

**INFORMATION AND THE CHANGE IN THE PARADIGM IN ECONOMICS,  
PART 1**by **Joseph E. Stiglitz\***

The research for which George Akerlof, Mike Spence, and I are being recognized is part of a larger research program which, today, embraces hundred, perhaps thousands, of researchers around the world. In this lecture, I want to set the particular work which was cited within this broader agenda, and that agenda within the broader perspective of the history of economic thought. I hope to show that Information Economics represents a fundamental change in the prevailing paradigm within economics. Problems of information are central to understanding not only *market* economics but also *political economy*, and in the last section of this lecture, I explore some of the implications of information imperfections for political processes.

**INTRODUCTION**

Many years ago Keynes wrote:

The ideas of economists and political philosophers, both when they are right and when they are wrong, are more powerful than is commonly understood. Indeed, the world is ruled by little else. Practical men, who believe themselves quite exempt from any intellectual influences, are usually the slaves of some defunct economist. Madmen in authority, who hear voices in the air, are distilling their frenzy from some academic scribbler of a few years back. Keynes [1936].

Information economics has already had a profound effect on how we think about economic policy, and is likely to have an even greater influence in the future. The world is, of course, more complicated than our simple—or even our more complicated models—would suggest. Many of the major political debates over the past two decades have centered around one key issue: the efficiency of the market economy, and the appropriate relationship between the market and the government. The argument of Adam Smith [1776], the founder of modern economics, that free markets led to efficient outcomes, “as if by an invisible hand” has played a central role in these debates: it suggested that we could, by and large, rely on markets *without government interven-*

*tion*. There was, at best, a limited role for government. The set of ideas that I will present here undermine Smith’s theory and the view of government that rested on it. They have suggested that the reason that the hand may be invisible is that it is simply not there—or at least that if it is there, it is palsied.

When I began the study of economics some forty one years ago, I was struck by the incongruity between the models that I was taught and the world that I had seen growing up, in Gary Indiana, a city whose rise and fall paralleled the rise and fall of the industrial economy. Founded in 1906 by U.S. Steel, and named after its Chairman of the Board, it had declined to but a shadow of its former self by the end of the century. But even in its heyday, it was marred by poverty, periodic unemployment, and massive racial discrimination. Yet the theories that we were taught paid little attention to poverty, said that all markets cleared—including the labor market, so unemployment must be nothing more than a phantasm, and that the profit motive ensured that there could not be economic discrimination.<sup>1</sup> If the central theorems that argued that the economy was Pareto efficient—that, in some sense, we were living in the best of all possible worlds—were true, it seemed to me that we should be striving to create a different world. As a graduate student, I set out to try to create models with assumptions—and conclusions—closer to those that accorded with the world I saw, with all of its imperfections.

\* 2001 Nobel Laureate, Columbia Business School, Columbia University, 1022 International Affairs, Building, 420 West 118th Street, New York, NY 10027, USA.

Due to the length of the essay, the first part is published in the Fall issue, and the second part, which includes all the notes and references, will be published in the Spring 2004 issue. A complete set of notes and references is also available by contacting Richard Nilsen at e-mail address rn71238@pace.edu

My first visits to the developing world in 1967, and a more extensive stay in Kenya in 1969, left an indelible impression on me. Models of perfect markets, as badly flawed as they might seem for Europe or America, seemed truly inappropriate for these countries. But while many of the key assumptions that went into the competitive equilibrium model seemed not to fit these economies well, the ones that attracted my attention were the imperfection of information, the absence of markets, and the pervasiveness and persistence of seeming dysfunctional institutions, like sharecropping. With workers having to surrender 50% or more of their income to landlords, surely (if conventional economics were correct), incentives were greatly attenuated. Traditional economics said not only that institutions (like sharecropping)<sup>2</sup> did not matter, but neither did the distribution of wealth. But if workers owned their own land, then they would not face what amounted to a 50% tax. Surely, the distribution of wealth did matter.

I had seen cyclical unemployment—sometimes quite large—and the hardship it brought as I grew up, but I had not seen the massive unemployment that characterized African cities, unemployment that could not be explained either by unions or minimum wage laws (which, even when they existed, were regularly circumvented). Again, there was a massive discrepancy between the models we had been taught and what I saw.

The new ideas and models were not only useful in addressing broad philosophical questions, such as the appropriate role of the state, but also in analyzing concrete policy issues. In the 1970s, economists became increasingly critical of traditional Keynesian ideas, partly because of their assumed lack of micro-foundations. The attempts made to construct a new macro-economics based on *traditional* micro-economics, with its assumptions of well functioning markets, were doomed to failure. Recessions and depressions, accompanied by massive unemployment, were symptomatic of massive market failures. The market for labor was clearly not clearing. How could a theory that began with the assumption that all markets clear ever provide an explanation? If individuals could easily smooth their consumption by borrowing at safe rates of interest, then the relatively slight loss of lifetime income caused by an interruption of work of six months or a year would hardly be a problem; but the unemployed do not have access to capital markets,

at least not at reasonable terms, and thus unemployment is a cause of enormous stress. If markets were perfect, individuals could buy private insurance against these risks; yet it is obvious that they cannot. Thus, one of the main developments to follow from this line of research into the consequences of information imperfections for the functioning of markets is the construction of *macro economic* models that help explain why the economy amplifies shocks and makes them persistent, and why there may be, even in competitive equilibrium, unemployment and credit rationing.

I believe that some of the huge mistakes which have been made in policy in the last decade, in for instance the management of the East Asia crisis or the transition of the former communist countries to the market, might have been avoided if there had been a better understanding of issues, like bankruptcy and corporate governance, to which the new information economics called attention. And the so-called Washington consensus policies<sup>3</sup>, which have predominated in the policy advice of the international financial institutions over the past quarter century, have been based on market fundamentalist policies which ignored the information-theoretic concerns, and this explains at least in part their widespread failures.<sup>4</sup>

Information affects decision making in every context—not just inside firms and households. More recently, I have turned my attention to some aspects of what might be called the *political economy* of information: the role of information in political processes, in collective decision making. For two hundred years, well before the economics of information became a subdiscipline within economics, Sweden had enacted legislation to increase transparency. There are asymmetries of information between those governing and those governed, and just as markets strives to overcome asymmetries of information, we need to look for ways by which the scope for asymmetries of information in political processes can be limited and their consequences mitigated.

## THE HISTORICAL SETTING

I do not want here to review and describe in detail the models of information asymmetries that have been constructed in the past thirty years. In recent years, there have been a number of survey

articles,<sup>5</sup> even several books,<sup>6</sup> and interpretative essays.<sup>7</sup> I do want to highlight some of the dramatic impacts that information economics has had on how economics is approached today, how it has provided explanations to phenomena that were previously unexplained, how it has altered our views about how the economy functions, and, perhaps most importantly, how it has led to a rethinking of the appropriate role for government in our society. In describing the ideas, I want to trace out some of their origins: to a large extent, they were responses to attempts to answer specific policy questions or to explain specific phenomena to which the standard theory provided an inadequate explanation. But any discipline has a life of its own, a prevailing paradigm, with assumptions and conventions. Much of the work was motivated by an attempt to explore the limits of that paradigm—to see how the standard models could embrace problems of information imperfections (which turned out not to be very well.)

For more than a hundred years, formal modeling in economics has focused on models in which information was perfect. Of course, everyone recognized that information was in fact imperfect, but the hope, following Marshall's dictum "Natura non facit saltum," was that economies in which information was not too imperfect would look very much like economies in which information was perfect. One of the main results of our research showed that this was not true; that even a small amount of information imperfection could have a profound effect on the nature of the equilibrium.

The reigning paradigm of the twentieth century, the neoclassical model, ignored the warnings of the nineteenth century and earlier masters on how information concerns might alter the analyses, perhaps because they could not see how to embrace them in their seemingly precise models, perhaps because doing so would have led to uncomfortable conclusions about the efficiency of markets. For instance, Smith, in anticipating later discussions of adverse selection, wrote that as firms raise interest rates, the best borrowers drop out of the market. *If* lenders know perfectly the risks associated with each borrower, this would matter little; each borrower would be charged an appropriate risk premium. It is because lenders do not know the default probabilities of borrowers perfectly that this process of adverse selection has such important consequences.<sup>8</sup>

I have already noted in the introduction that *something* was wrong—seriously wrong—with the competitive equilibrium models that represented the prevailing paradigm when we went to graduate school. It seemed to say that unemployment didn't exist, that issues of efficiency and equity could be neatly separated, so that economists could neatly set aside problems of inequality and poverty as they went about their business of designing more efficient economic systems. But there were a host of other predictions, empirical puzzles, that were hard to reconcile with the standard theory: in micro-economics, there were tax paradoxes such as why did firms seemingly not take actions which minimized their tax liabilities, security market paradoxes, such as why did asset prices seem to exhibit such high volatility,<sup>9</sup> and behavioral puzzles, such as why did firms respond to risks in ways which were markedly different from that predicted by the theory.<sup>10</sup> In macro-economics, the cyclical movements of many of the key aggregate variables, such as consumption,<sup>11</sup> inventories,<sup>12</sup> real product wages,<sup>13</sup> real consumption wages,<sup>14</sup> and interest rates<sup>15</sup> are hard to reconcile with the standard theory, and if the perfect market assumptions were even approximately satisfied, the distress caused by cyclical movements in the economy would be much less than seems to be the case.<sup>16</sup>

The problems that we saw with the models that we were taught was not only that they seemed wrong, but that they left a host of phenomena and institutions unexplained—why were IPO's typically sold at a discount? Why did equities, which provided far better risk diversification than debt, play such a limited role in financing new investment?<sup>17</sup>

There were, to be sure, some Ptolemaic attempts to defend and elaborate on the old model. Some, like George Stigler,<sup>18</sup> while recognizing the importance of information, argued that once the real costs of information were taken into account, even with imperfect information, the standard results of economics would still hold. Information was just a transactions cost. In the approach of many Chicago economists, information economics was like any other branch of applied economics; one simply analyzed the special factors determining the demand and supply for information, just as agricultural economics analyzed those factors affecting the market for wheat. For the more mathematically inclined, information could be incorporated into production functions of, say, goods by inserting an "I" for the

input “information,” and “I” itself could be produced by inputs, like labor. Our analysis showed that that this approach was wrong, as were the conclusions derived from it.

Practical economists who could not ignore the bouts of unemployment which had plagued capitalism since its inception talked of the neoclassical synthesis: using Keynesian interventions to ensure that the economy remained at full employment, and once that was done, the standard neoclassical propositions would once again be true. But while the neoclassical synthesis<sup>19</sup> had enormous intellectual influence, by the 1970s and 80s it came under attack from two sides. It was an assertion, not based on a coherent view of the economy. One side attacked the underpinnings of Keynesian economics, its micro-foundations; why would rational actors be out of equilibrium—with unemployment persisting—in the way that Keynes had suggested. This side effectively denied the phenomena which Keynes was attempting to explain.

Worse still, some saw unemployment as largely reflecting an interference (e.g. by government in setting minimum wages, or trade unions, in using their monopoly power to set wages too high) with the free workings of the market, with the obvious implication: unemployment would be eliminated if markets were made more *flexible*, that is, if unions and government interventions were eliminated. Even if wages fell a third in the Great Depression, they should have, in this view, fallen even more.

There was an alternative perspective (articulated more fully in Greenwald and Stiglitz, 1987a, 1988b): why shouldn't we believe that massive unemployment was just the tip of the iceberg, of more pervasive market efficiencies that were harder to detect. If markets seemed to function *so* badly some of the time, certainly they must be malperforming in more subtle ways much of the time. The economics of information bolstered the latter view.

Similarly, given the nature of the debt contracts, the falling prices led to bankruptcy and economic disruptions, actually exacerbating the economic downturn. Had there been more wage and price flexibility, matters might have been even worse. Moreover, neither government nor unions imposed the limitations on wage and price dynamics in many sectors of the economy; at the very least, those who argued that the problem was wage and price rigidities had to look for other *market* imperfections, and

any policy remedy (including a call for greater flexibility) had to take those factors into account.

In the next section, I shall explain how it was not just the discrepancies between the standard competitive model and its predictions which lead to its being questioned. The model was not robust—even slight departures from the underlying assumption of perfect information had large consequences.

But before turning to those issues, it may be useful to describe some of the concrete issues which underlay the beginnings of my research program in this area. Key in my thinking on these issues was the time between 1969 and 1971 that I spent at the Institute for Development Studies at the University of Nairobi with the support of the Rockefeller Foundation.

### *Education as a screening device*<sup>20</sup>

The newly independent Kenyan government, as it attempted to forge policies which would promote their growth and development, was asking questions that never seemingly been raised by their colonial masters. How much should it invest in education? It was clear that a better education got one better jobs—the credential put one at the head of the job queue. Gary Fields, a young scholar working at the Institute of Development Studies there, developed a simple model<sup>21</sup> suggesting that the private returns to education—the enhanced probability of getting a good job—differed from the social return; and that it was possible that as more people get educated, the private returns got higher (it was even more necessary to get the credential) even though the social return might decline. Here, education was performing a markedly different function than it did in traditional economics literature, where it simply added to human capital and improved productivity.<sup>22, 23</sup> The analysis had important implications for Kenya's decision about how much to invest in higher education. The problem with Fields' work was that it did not provide a full *equilibrium* analysis: wages were fixed, rather than competitively determined.

This led me to ask, what would the market equilibrium look like if wages were set equal to mean marginal products *conditional on the information that was available*? And this in turn forced me to ask: what were the *incentives* and *mechanisms* for employers and employees to acquire or transmit information? Within a group of otherwise similar

job applicants (who therefore face the same wage), the employer has an incentive to identify who is the most able, to find some way of *sorting* or *screening* among them, *if he could keep that information private*. But he often can't, and if others find out about the true ability, the wage will be bid up, and he will be unable to *appropriate* the return to the information. At the very beginning of this research program we had thus identified one of the key issues in information economics, the difficulty of *appropriating* the returns.

On the other hand, the employee, *if he knew his ability* (that is, if there were *asymmetries of information between the employee and the employer*) and he knew that his abilities were above the average of those in the market, he would have an incentive to convince the employer of his ability. But someone at the bottom of the ability distribution had an incentive not have the information revealed. Here was a second principle that was to be explored in subsequent years: there are incentives on the part of individuals for information not to be revealed, for secrecy, or, in modern parlance, for a lack of transparency. This raised a question: how did the forces for secrecy and for information disclosure get balanced? What was the equilibrium that emerged? I will postpone until the next section a description of that equilibrium.

### *Efficiency wage theory*

That summer in Kenya I began three other research projects related to information imperfections. At the time I was working in Kenya, there was heavy urban unemployment. My colleagues at the Institute for Development Studies, Michael Todaro and John Harris had formulated a simple model of labor migration from the rural to the urban sector which accounted for the unemployment. High urban wages attracted workers, and they were willing to risk unemployment for the chance of those higher wages.<sup>24</sup> Here was a simple, general equilibrium model of unemployment, but again there was one missing piece: how could you explain the high wages, which were well in excess of the minimum wage? It did not seem as if either government or unions were *forcing* these high wages. One needed an equilibrium theory of wage determination. I recalled, during an earlier stint at Cambridge, discussions with Harvey Leibenstein who had postulated that in very poor countries, higher

wages lead to higher productivity.<sup>25</sup> It might not pay firms to cut wages, if productivity was cut more than proportionately, even if there was an excess supply of labor. The key insight was to recognize that there were a variety of other reasons why, when information and contracting were imperfect, productivity might depend on wages.<sup>26</sup> In that case, it might pay firms to pay a higher wage than the minimum necessary to hire labor; such wages I referred to as *efficiency wages*. With efficiency wages, there could exist an equilibrium level of unemployment. I explored four explanations for why productivity might depend on wages (besides nutrition). The simplest was that lower wages lead to higher turnover, and therefore the higher turnover costs which the firm bore.<sup>27</sup> It was not until some years later that we were able to explain more fully—based on limitations of information—why it was that firms had to bear these turnover costs.<sup>28</sup>

But there was another version of the efficiency wage related to the work I was beginning on asymmetric information. Any manager will tell you that you attract better workers by paying them higher wages. This was just an application of the general notion of adverse selection, which played a central role in earlier insurance literature, where firms had long recognized that as they charge a higher premium, the best risks stopped buying insurance.<sup>29</sup> Firms in a market do not passively have to accept the “market wage.” Even in competitive markets, firms could, if they wanted, offer higher wages than others. Market clearing was not a constraint on firms. If all firms were paying the market-clearing wage, it might pay a firm to offer a higher wage, to attract more able workers. The efficiency wage theory meant that there could exist unemployment *in equilibrium*.

It was thus clear that the notion that had underlay much of traditional competitive equilibrium analysis—that markets *had* to clear—was simply not true if information were imperfect.

The formulation of the efficiency wage theory that has received the most attention over the years, however, has been the one that has focused on problems of *incentives*. Many firms claim that paying high wages induces their workers to work harder. The problem that Carl Shapiro and I [1984] faced was to try to make sense of this claim. If all workers are identical, and were paid the same wage, then if it paid one firm to pay a high wage, it would pay all of them. But if a worker was then fired for shirk-

ing, and there was full employment, he could immediately get another job, at the same wage. The high wage would provide no incentive. But if there was unemployment, then there was a price for shirking. We showed that *in equilibrium* there had to be unemployment: unemployment was the discipline device that forced workers to work.<sup>30</sup> The model had strong policy implications, some of which I shall describe below. Our work illustrated the use of highly simplified models to help clarify thinking about quite complicated matters. In practice, of course, workers are not identical, so problems of adverse selection become intertwined with those of incentives; being fired does convey information—there is typically a stigma.

(There was a fourth version of the efficiency wage, where productivity was related to *morale* effects, perceptions about how *fairly* they were being treated. While I briefly discussed this version in my earlier work,<sup>31</sup> it was not until almost twenty years later that the idea was fully developed, in the important work of Akerlof and Yellen [1990].)

### *Sharecropping and the general theory of incentives*

This work on the economics of incentives in labor markets was closely related to the third research project that I began in Kenya. In traditional economic theory, while considerable lip service was paid to incentives, there was no real incentive issue. With perfect information, individuals are paid to perform a particular service; if they perform it they receive the amount contracted for; if they do not perform, they receive nothing. With imperfect information, firms have to motivate and monitor, rewarding them for observed good performance and punishing them for bad. My interest in the issues was first aroused by thinking about sharecropping, a common form of land tenancy in a developing country, where the worker surrenders half (sometimes two thirds) of the produce to the landlord in return for the use of his land. At first blush, this seemed a highly inefficient arrangement; it was equivalent to a 50% tax on workers' labor. But what were the alternatives? The worker could rent the land, but that meant he had to bear all the risk of fluctuations in output; and besides, he often did not have the requisite capital. He could work as wage labor, but that meant that the landlord would have to monitor him, to ensure that he worked. Sharecrop-

ping represented a compromise, between risk bearing and incentives. The underlying *information* problem was that the input of the worker could not be observed, but only his output, and his output was not perfectly correlated with his input. The sharecropping contract could be thought of as a combination of a rental contract *plus* an insurance contract, in which the landlord "rebates" part of the rent if crops turn out badly. There is not full insurance (which would be equivalent to a wage contract) because such insurance would attenuate all incentives. The adverse effect of insurance on incentives to avoid the insured against contingency is referred to as *moral hazard*.<sup>32</sup> In Stiglitz [1974b] I analyzed the *equilibrium* sharecropping contract. In that paper, I recognized that the incentive problems I explored there were isomorphic to those facing modern corporations, e.g. in providing incentives to their managers,<sup>33</sup> and there followed a large literature on optimal and equilibrium incentive schemes<sup>34</sup> in labor, capital, and insurance markets. Contracts had to be based on observables, like processes (or which crops were grown) and *observable* inputs (like fertilizers). Many of the results obtained earlier in the work on adverse selection had their parallel in this area of "adverse incentives."<sup>35</sup> For instance, with Richard Arnott I analyzed the equilibrium [1988a, 1988b], which entails partial insurance.

### *Equilibrium wage and price distributions*

The fourth strand of research looked at the issue of wage differentials that I had observed from a different perspective. The work on labor turnover had suggested that firms that faced higher turnover costs might pay higher wages. But one of the reasons that individuals quit was to obtain a higher paying job. The turnover rate depended on the wage distribution. The challenge was to formulate an *equilibrium* model, in which there was a wage distribution, which led firms to charge different wages—the distribution of wages that had originally been postulated.

More generally, efficiency wage theory said that it paid firms to pay a higher wage than necessary to obtain workers; but the level of the efficiency wage could vary across firms; for instance, firms with higher turnover costs, or where worker inefficiency could lead to large losses of capital, or where monitoring was more difficult, might find it desirable to

pay higher wages. The implication was that similar labor might receive quite different compensation; wage discrepancies might not be explicable solely in terms of differences in abilities.

I was to return to these four themes repeatedly in my research over the following three decades.

## FROM THE COMPETITIVE EQUILIBRIUM PARADIGM TO THE INFORMATION PARADIGM

In the previous section, I described how my experiences, especially in Kenya—the disparities between the models used and the world that I saw—had motivated a search for an alternative paradigm. But there was another motivation, driven more by the internal logic and structure of the competitive model itself, which was the dominant paradigm thirty years ago.

The model virtually made economics a branch of engineering (with no aspersions to that noble profession), and the participants in the economy better or worse engineers. Each was solving a maximization problem, with full information: households maximizing utility subject to budget constraints, firms maximizing profits (market value), and the two interacting in competitive product, labor, and capital markets. One of the peculiar implications was that there never were disagreements about what the firm should do: alternative management teams would presumably come up with the same solution to the maximization problems. Another peculiar implication was the meaning of risk: when a firm said that a project was risky, that (should have) meant that it was highly correlated with the business cycle, not that it had a high chance of failure.<sup>36</sup> I have already described some of the other peculiar implications of the model: the fact that there was no unemployment or credit rationing, that it focused on only a limited subset of the information problems facing society, that it seemed not to address key issues—like incentives and motivation.

But much of the research in the profession is directed not at these big lacunae, but at seemingly more technical issues—at the mathematical structures. The underlying *mathematics* required assumptions of convexity and continuity, and with these assumptions one could prove the *existence* of equilibrium and its (Pareto) efficiency. The standard proofs of these fundamental theorems of wel-

fare economics did not even list in their enumerated assumptions those concerning information: the perfect information assumption was so ingrained it did not have to be explicitly stated. The *economic* assumptions to which the proofs of efficiency called attention concerned the absence of externalities and public goods. The market failures approach to the economics of the public sector<sup>37</sup> discussed alternative approaches by which these market failures could be corrected, but these market failures were highly circumscribed.

There was, moreover, a curious disjunction between the language economists used to explain markets and the models they constructed. They talked about the information efficiency of the market economy, though they focused on a single information problem, that of scarcity. But there are a myriad of other information problems faced by consumers and firms every day, concerning for instance the prices and qualities of the various objects that are for sale in the market, the quality and efforts of the workers they hire, the returns of investment projects. In the standard paradigm, the competitive general equilibrium model,<sup>38</sup> there were no shocks, no unanticipated events: at the beginning of time, the full equilibrium was solved, and everything from then on was an unfolding over time of what had been planned in each of the contingencies. In the real world, the critical question was how, and how well, do markets handle these information problems?

There were other aspects of the standard paradigm which seemed hard to accept. It argued that institutions did not matter—markets could see through them, and equilibrium was simply determined by the laws of supply and demand. It said that the distribution of wealth did not matter.<sup>39</sup> And it said that (by and large) history did not matter—knowing preferences and technology and initial endowments, one could describe the time path of the economy.<sup>40</sup>

Work on the economics of information began by questioning each of the underlying premises, each of the central theorems. Consider, to begin with, the mathematical structures that had underlay some much of the formalization of economics of the latter half of the twentieth century, the convexity assumptions which corresponded to long standing principles of diminishing returns. With imperfect information (and the costs of acquiring it) these assumptions were no longer plausible. It was not

just that the cost of acquiring information could be viewed as fixed costs.<sup>41</sup> Work with Roy Radner (Radner and Stiglitz [1984]) showed that there was a *fundamental non-concavity in the value of information*, that is, under quite general conditions, it never paid to buy just a little bit of information. Work with Richard Arnott (Arnott and Stiglitz [1988a]) showed that such problems were pervasive in even the simplest of moral hazard problems (where individuals had a choice of alternative actions, e.g. the amount of risk to undertake.) While we had not repealed the law of diminishing returns, we had shown its domain to be more limited than had previously been realized.<sup>42</sup>

Michael Rothschild and I showed that under natural formulations of what might be meant by a competitive market with imperfect information, equilibrium often did not exist<sup>43</sup>—even when there was an arbitrarily small amount of information imperfection.<sup>44</sup> While subsequent research has looked for alternative definitions of equilibrium,<sup>45</sup> we remain unconvinced; most of them violate the natural meaning of competition, i.e. where each participant in the market is so small that he believes that he will have no effect on the behavior of others. (Rothschild and Stiglitz [1997]).

The new information paradigm went further in undermining the foundations of competitive equilibrium analysis, the basic “laws” of economics, which include: the law of demand and supply (holding that market equilibrium was characterized by market clearing), the law of the single price, holding that the same good sold for a single price throughout the market, the law of the competitive price, holding that in equilibrium price equaled marginal cost, the efficient markets hypothesis, holding that in stock markets prices convey all the relevant information from the informed to the uninformed. Each of these cornerstones was rejected, or was shown to hold under much more restrictive conditions.

- We have shown how, when prices affect “quality”—either because of incentive or selection effects—equilibrium may be characterized by demand not equaling supply; firms will not pay lower wages to workers, even when they can obtain such workers, because doing so will raise their labor costs; firms will not charge higher interest rates, even when they can do so, because of an excess demand for credit,

because doing so will increase the average default rate, and thus lower expected returns.

- We have shown that the market will be characterized by wage and price distributions, even when there is no exogenous source of “noise” in the economy, even when all firms and workers are (otherwise) identical.
- We have shown that in equilibrium, firms will charge a price in excess of the marginal costs, or workers are paid a wage in excess of their reservation wage. The “surplus” is required to provide the *incentive* for maintaining a reputation.<sup>46</sup> Even in situations where reputation rents were not required, information imperfections gave rise to market power—there is *imperfect competition*—which results in firms charging prices in excess of marginal cost.<sup>47</sup>
- The efficient markets hypothesis<sup>48</sup> held that prices in the stock market fully reflected all information. But if that were the case, then there would be no incentive for anyone to expend money to collect information. Work with Sanford Grossman [1976, 1980a] showed that the price system both imperfectly aggregated information and that there was an equilibrium amount of “disequilibrium.”

The most fundamental reason why markets with imperfect information differ from the perfect ones is that actions (including choices) convey information, market participants know this, and this affects their behavior.

A decision by a firm to provide a guarantee is not just a matter that the firm is better able to absorb the risk of a product failure; his willingness to provide a guarantee conveys information about his confidence in the product. An insured is willing to take a policy with a large deductible not because he is not risk averse, but because this action conveys information to the insurance company, that he is willing to bear the risk because he thinks the likelihood of an accident is low. At the same time, a firm may not assign an employee to a highly visible job because it knows that the assignment will be interpreted as an indication that the employee is good, making it more likely that a rival will try to hire the person away. Even if he fails, the current employer will have to pay a higher salary.

One of the early insights (Akerlof, 1971) was that markets may be thin or absent. One of the standard assumptions of the old paradigm was that there was a complete set of markets—including intertem-



poral markets (capital markets) and risk markets. The absence of particular markets, e.g. for risk, has profound implications for how *other* markets function. The fact that workers and firms cannot buy insurance against many of the risks which they face affects labor and capital markets; it leads, for instance, to labor contracts in which the employer provides *some* insurance. But the design of these more complicated, but still imperfect and incomplete, contracts, affects the efficiency, and overall performance, of the economy.

Perhaps most importantly, under the standard paradigm, markets are Pareto efficient, except when one of a limited number of market failures occurs. Under the imperfect information paradigm, markets are almost never Pareto efficient.

While information economics thus undermined these long standing principles of economics, it also provided explanations for many phenomena that had long been unexplained. Earlier, I mentioned the seemingly inefficient institution of sharecropping, for which information economics provided an explanation. Before turning to these *applications*, I want to present a somewhat more systematic account of the *principles* of the economics of information.

### *Some problems in constructing an alternative paradigm*

The fact that information was imperfect was, of course, well recognized by all economists. While they may have *hoped* that economies with imperfect information behaved much like economies with perfect information, the real reason that models with imperfect information were not developed was that it was not obvious how to do so. There were several problems that had to be overcome: while there was a single way in which information is perfect, there are an infinite number of ways in which information can be imperfect. One of the keys to success was formulating simple models in which the set of relevant information could be fully specified—and so the precise ways in which information was imperfect could also be fully specified. But there was a danger in this methodology, as useful as it was: in these over simplistic models, there were sometimes ways in which there could be full information revelation; the information problems could be fully resolved. In the real world, of course, this never happens, which is why in some of the later

work (e.g. Grossman and Stiglitz [1976, 1980a], we worked with models with an infinite number of states.<sup>49</sup>

Perhaps the hardest problem was modeling equilibrium. It was important to think about both sides of the market—employers and employees, the insurance company and the insured, lender and borrower. Each had to be modeled as “rational,” in some sense, making inferences on the basis of available information. Each side’s behavior too had to be rational, based on beliefs about the consequences of their actions; and those consequences in turn depended on what inferences others would draw from those actions. I wanted to model *competitive* behavior, where each actor in the economy was small, and believed he was small—and so his actions could not or would not affect the equilibrium (though others’ inferences about himself might be affected). Finally, one had to think carefully about what was the feasible set of actions: what might each side do to extract or convey information to others.

As we shall see, the variety of results obtained (and much of the confusion in the early literature) arose partly from a failure to be as clear as one might about the assumptions. For instance, the standard adverse selection model had the quality of the good offered in the market (say of used cars, or riskiness of the insured) depending on price. The car buyer (the seller of insurance) knows the *statistical* relationship between price and quality, and this affects his demand. The market equilibrium is the price at which demand equals supply. But that is an equilibrium if and only if there is no way by which the seller of a good car can convey that information to the buyer—so that he can earn a quality premium—and if there is no way by which the buyer can sort out good cars from bad cars. Typically, there are such ways, and it is the attempt to elicit that information which has profound effects on how markets function. To develop a new paradigm, we had to break out from long established premises, to ask what should be taken as *assumptions* and what should be derived from the analysis. Market clearing could not be taken as an assumption; neither could the premise that a firm sells a good at a particular price to all customers. One could not *begin* the analysis even by assuming that in competitive equilibrium there would be zero profits. In the standard theory, if there were positive profits, a firm might enter, bidding away existing customers. In

the new theory, the attempt to bid away new customers by slightly lowering prices might lead to marked changes in their behavior or in the mix of customers, in such a way that the profits of the new entrant actually became negative. One had to rethink all the conclusions from first premises.

We made progress in our analyses because we began with highly simplified models of *particular* markets, that allowed us to think through carefully each of the assumptions and conclusions. From the analysis of particular markets (whether the insurance market, as in Rothschild Stiglitz, the education market, the labor market, or the land tenancy/sharecropping market), we attempted to identify general principles, to explore how these principles operated in each of the other markets. In doing so, we identified particular features, particular *informational assumptions*, which seemed to be more relevant in one market or another. The nature of competition in the labor market is different than that in the insurance market or the capital market, though they have much in common. This interplay, between looking at the ways in which such markets are similar and dissimilar, proved to be a fruitful research strategy.<sup>50</sup>

## SOURCES OF ASYMMETRIES OF INFORMATION

Information imperfections are pervasive in the economy: indeed, it is hard to imagine what a world with perfect information would be like. Much of the research I will describe below focuses on *asymmetries* of information, the fact that different people know different things: workers know more about their ability than does the firm; the person buying insurance knows more about his health, whether he smokes and drinks immoderately, than the insurance firm; the owner of a car knows more about the car than potential buyers; the owner of a firm knows more about the firm than a potential investor; the borrower knows more about his risk and risk taking than the lender.

The essential feature of a decentralized market economy is that different people know different things; in this sense, economists had long been thinking of markets with information asymmetries. But the earlier literature had neither thought about how they were created, nor what their consequences might be. Moreover, while much of the

earlier literature focused on simple situations of information asymmetry—such as those described in the previous paragraphs, the problems of information *imperfections* run deeper, and the research described below discusses some of these more general results. The individual may know little about his true health condition; the insurance company, through a simple examination, might even become more informed (at least concerning relevant aspects, e.g. implications for life expectancy).

Some of these information asymmetries are inherent: the individual naturally knows more about himself than does anyone else. Some of the asymmetries arise naturally out of economic processes. The current employer knows more about the employee than other potential employers; a firm may find out a great deal of information in the process of dealing with suppliers that others may not know; the owner of a car naturally knows the faults of the car better than others—and in particular, he knows whether or not he has a lemon. While such information asymmetries inevitably arise, the extent to which they do so and their consequences depend on how the market is structured, and the recognition that they will arise affects market behavior. For instance, one of the important insights of work in this area is to show how information asymmetries lead to thin or non-existent markets (Akerlof [1970]). But this means that even if an individual has no more information about his ability than potential employers, the moment he goes to work for an employer, an information asymmetry has been created—the employer may know more about the individual's ability than others. The consequence is that the “used labor” market does not work well. Others will be more tame in bidding for his services, knowing that they will succeed in luring him away from his current employer only if they bid too much. If they bid less than his productivity, his current employer will match. Labor mobility is impeded. But that gives market power to the first employer, which he will be tempted to exercise. The recognition of this naturally affects even the “new labor” market. Because an individual is locked into a job, he will be more risk averse in accepting an offer. The terms of the initial contract have to be designed to reflect the diminution of the workers' bargaining power and his reduced labor mobility that occurs immediately after signing.<sup>51, 52</sup>

## OVERCOMING INFORMATION ASYMMETRIES

I now want to discuss briefly the ways by which information asymmetries are dealt with, how they can be (partially) overcome.

### *Incentives for gathering and disclosing information*

There are two key issues: what are the *incentives* for obtaining information, and what are the *mechanisms*. My brief discussion of the analysis of education as a screening device suggested the fundamental incentives: more able individuals (lower risk individuals, firms with better products) will receive a higher wage (will have to pay a lower premium, will receive a higher price for their products) if they can establish that they are more productive (lower risk, higher quality).

We noted earlier that while some individuals have an incentive to disclose information, those who are less able have an incentive not to have the information disclosed. Was it possible that in market equilibrium, only *some* of the information would be revealed? One of the early important results was that, if the more able succeed (costlessly) to establish that they are more able, then the market will be fully revealing, even though all of those who are below average would prefer that no information be revealed. In the simplest models, I described a process of unraveling: if the most able could establish his ability, he would; but then all but the most able would be grouped together, receiving the mean marginal product of that group; and the most able of that group would have an incentive to reveal his ability. And so on down the line, until there was full revelation.<sup>57</sup>

What happens if those who are more able cannot credibly convince potential employers of their ability (or if those who are low risk cannot convince potential insurance companies)? The other side of the market has an incentive too to gather information: an employer that can find a worker who is better than is recognized by others will have found a bargain; his wage will be determined by what others think of him. The problem, as we noted, is that if what he knows becomes known to others, the wage will be bid up, and he will be unable to appropriate the returns on his investment in information acquisition.

To take another example, it is natural that in the process of oil exploration, a company finds out information that is relevant for the likelihood that there will be oil in a neighboring tract. There is an informational externality.<sup>53</sup> The existence of this asymmetric information affects the nature of the bidding for oil rights on the neighboring tract. Bidding where there is known to be asymmetries of information will be markedly different from that where such asymmetries do not exist.<sup>54</sup> Those who are uninformed will presume that they will win only if they bid too much—information asymmetries exacerbate the problem of the winners' curse.<sup>55</sup> The government (or other owners of large tracts to be developed) should take this into account in its leasing strategy. And the bidders in the initial leases too will take this into account: part of the value of winning in the initial auction is the information rent that will accrue in later rounds.

### *Creating asymmetries and imperfections of information*

While early work in the economics of information dealt with how markets overcame problems of information asymmetries, and information imperfections more generally, later work turned to how markets *create* information problems, partly in an attempt to exploit market power. Managers of firms attempt to *entrench* themselves, increasing their bargaining power, e.g. vis-à-vis alternative management teams,<sup>56</sup> and one of the ways that they do this is to take actions which *increase* information asymmetries. (Edlin and Stiglitz [1995]). Doing so effectively reduces competition in the market for management. This is an example of the general problem of corporate governance, to which I alluded earlier, and to which I will return later.

Similarly, the presence of information imperfections give rise to market power; and firms can exploit this market power through "sales" and other ways of differentiating among individuals who have different search costs. (Salop, 1977, Salop and Stiglitz, 1976, 1982, Stiglitz 1979a). The price dispersions which exist in the market are *created* by the market—they are not just the failure of markets to arbitrage fully price differences *caused* by shocks that affect different markets differently.

The fact that if there was competition, it would be difficult for the *screener* to appropriate the returns had an important implication: in markets where, for one reason or another, the more able (the firms with the better investment projects, the more able workers) cannot (fully) convey their attributes, *if there is to be investment in screening there must be imperfect competition in screening*. The economy, in effect, has to choose between two different imperfections: imperfections of information or imperfections of competition. Of course, in the end, there will be both forms of imperfection.<sup>58, 59</sup>

This is but one of many examples of the *interplay* between market imperfections. Earlier, for instance, we discussed the incentive problems associated with sharecropping, which arise when workers do not own the land that they till. This problem could be overcome if individuals could borrow money, to buy their land. But capital market imperfections—limitations on the ability to borrow, which themselves arise from information imperfections—explain why this “solution” does not work.

There is another important consequence: if markets were fully informationally efficient—that is, if information disseminated instantaneously and perfectly throughout the economy—then no one would have any incentive to gather information, so long as there was any cost of doing so. That is why markets cannot be fully informationally efficient. (See Grossman and Stiglitz, 1976, 1980a.)

### *Mechanisms for elimination or reducing information asymmetries*

In simple models where individuals know their own ability, or the insured knows his own risk, or the borrower knows his own likelihood of repaying, there might seem an easy way to resolve the problem of information asymmetry: let each person tell his true characteristic. The underlying problem arose from the fact that individuals did not necessarily have the incentive to tell the truth. Assume employees knew their abilities. An employer might ask, what is your ability? The more able might answer honestly. As we have seen, the least able would have an incentive to lie, to say that he was more able than he was. Talk is cheap. There had to be some other ways by which information could be credibly conveyed.

### *Screening by examination*

The simplest way by which that could be done was an exam. As I constructed a simple competitive equilibrium model,<sup>60</sup> two further general principles became apparent: *the gains of the more able were largely at the expense of the less able*; by establishing that an individual is of higher ability, thereby leading, in equilibrium, to higher wages, he simultaneously establishes that others are of lower ability. The private returns to expenditures on education exceed the social returns. It was clear that there were important *externalities* associated with information, a theme which was to recur in later work.

But a more striking result emerged: there could exist multiple equilibria, one in which information was fully revealed (the market identified the high and low ability people) and the other of which it was not (called a pooling equilibrium). The pooling equilibrium Pareto dominated the equilibrium with full revelation. This work, done some thirty years ago, established two results of important policy, which remarkably have not been fully absorbed into policy discussions even today. First, markets do not provide appropriate incentives for information disclosure. There is, in principle a role for government. And secondly, expenditures on information may be too great.<sup>61</sup>

### *The simplest adverse selection model*

But much of the information firms glean about their employees, banks about their borrowers, insurance companies about their insured comes not from examinations but from making *inferences* based on their behavior. This is a common practice in life—but not in our economic models. As I have already noted, the early discussions of adverse selection in insurance markets recognized that as an insurance company raised its premiums, those who were least likely to have an accident decided not to purchase the insurance; the willingness to purchase insurance at a particular price conveyed information to the insurance company.<sup>62</sup> George Akerlof recognized that this phenomenon was far more general: the willingness to sell a used car, for instance, conveyed information about whether the car was or was not a lemon.

Bruce Greenwald [1979, 1986] took the idea one important step further, showing how adverse selection applied to labor and capital markets:<sup>63</sup> The will-

ingness of an employer not to match the bid of a competitor conveyed information about the current employer's judgment of that individual's ability; the willingness of insiders in a firm to sell stock at a particular price conveyed information about the insider's view of the price relative to the expected return. Akerlof's claim that the result of these information asymmetries was that markets would be thin or absent helped explain why labor and capital markets often did not function well. It provided part of the explanation for why firms raised so little of their funds through equity (Mayer, 1990). Stigler was wrong: imperfect information was not just like transactions costs.

The consequences go well beyond just an absent or missing market. Weak equity markets meant that risks could not be divested, leading firms to act in a risk averse manner, explaining some of what would otherwise seem to be anomalous aspects of firm behavior.<sup>64</sup> These capital market imperfections, in turn, played a central role in the macro-economic theories to be described below. We have already described how the labor market imperfections—the limited mobility of labor and the firm's market power that results—affects the labor market, both *before* the asymmetry of information is created in the process of hiring and *after*.

### *The simplest adverse incentive model*

In the adverse selection model, individuals differed. There was a single action which conveyed information: they either entered or did not enter the particular market. But information imperfections also relate to *what people do*. A worker can work harder, a borrower can undertake greater risk and the insured can undertake greater care. The employer would like to know how hard his employee works; if he could, he would specify that in the contract; the lender would like to know the actions which borrower will undertake; if he could, he would specify that in the contract. These asymmetries of information about *actions* are as important as the earlier discussed asymmetries. Just as in the adverse selection model, the seller of insurance may try to overcome the problems posed by information asymmetries by *examination*, so too in the adverse incentive model, he may try to *monitor* the actions of the insured. But examinations and monitoring are costly and, while they yield some information, typically there remains a high level of residual infor-

mation imperfection. Just as in the adverse selection model, the seller of insurance recognizes that the average riskiness of the insurance applicants is affected by the terms of the insurance contract, so too the level of risk taking can be affected. And similar results hold in other markets. Borrowers' risk taking is affected by the interest rate charged. (Stiglitz and Weiss [1981]).

### *Efficiency wage theory, credit rationing*

While the early work in adverse selection explored the equilibrium in markets where the seller of insurance (the employer, the buyer of used cars, the lender) was rational enough to recognize the dependence of quality on price, he was not rational enough to exploit as fully as he could the information. While the law of supply and demand had been assumed to be a law of economics, there is in fact no law that requires the insurance firm to sell to all who apply at the premium he announces, the lender to lend to all who apply at the interest rate he announces, the employer to employ all those who apply at the wage he announces. So ingrained was the competitive equilibrium model in the mindset of economists that they simply assumed price-taking behavior. With perfect information and perfect competition, any firm that charged a price higher than the others would lose all of his customers; and at the going price, one faced a perfectly elastic supply of customers. In adverse selection and incentive models, what mattered was not just the supply of customers or employees or borrowers, but also their "quality"—the riskiness of the insured or the borrower, the returns on the investment, the productivity of the worker.

Since "quality" may increase with price, it may pay to offer a higher wage than the market clearing wage, for the lender to lend at an interest rate which exceeds the market clearing interest rate. This is true whether the dependence on quality arises from adverse selection or adverse incentive effects (or, in the labor market, because of morale or nutritional effects). And what matters is that there be imperfect information, not asymmetries of information. The healthy who decide not to buy insurance at a high premium do not need to *know* that they are healthy; they could be as uninformed as the insurance company, but simply—perhaps because of their health—have different preferences, e.g. they prefer to spend more of their money on recreational sports.

The consequence, as we have noted, is that market equilibrium may be characterized by demand not equaling supply: in equilibrium, the interest rate is lower than that at which the demand for loans equals the supply—there is credit rationing (Stiglitz and Weiss [1981], Keeton [1979]); the wage rate is higher than that at which the demand for labor equals the supply—there is unemployment.<sup>65</sup>

### *Conveying information through actions*

There is a much richer set of actions which convey information beyond those on which traditional adverse selection models have focused. An insurance company wants to attract healthy applicants. It might realize that by locating itself on the fifth floor of a walk up building, only those with a strong heart would apply. The willingness or ability to walk up five floors conveys information. More subtly, it might recognize that how far up it needs to locate itself, if it only wants to get healthy applicants, depends both on the premium charged and how high it locates itself. Or it may decide to throw in for free a membership in a health club, but charge a higher premium. Those who value a health club—because they will use it—willingly pay the higher premium. But these individuals are likely to be healthier.

There are a host of other actions which convey information. The quality of the guarantee offered by a firm can convey information about the quality of the product; only firms that believe that their product is reliable will be willing to offer a good guarantee. The guarantee is desirable not just because it reduces risk, but because it conveys information. The number of years of schooling may convey information about the ability of an individual. More able individuals may go to school longer, in which case the increase in wages associated with an increase in schooling may not be a consequence of the human capital that has been added, but rather simply be a result of the sorting that occurs.<sup>66</sup> The size of the deductible that an individual chooses in an insurance policy may convey information about his view about the likelihood of an accident or the size of the accidents he anticipates—*on average*, those who are less likely to have an accident may be more willing to accept high deductibles. The willingness of an entrepreneur to hold large fractions of his wealth in a firm (or to retain large fractions of the shares of the firm) conveys information about

his beliefs in the firm's future performance. If a firm promotes an individual to a particular job, it may convey information about the firm's assessment of his ability.

The fact that these actions may convey information affects behavior. In some cases, the action will be designed to obfuscate, to limit information disclosure. The firm that knows that others are looking at who it promotes, and that it will compete more vigorously for those, may affect the willingness of the firm to promote some individuals or assign them to particular jobs. (Waldman, 1984.) In others, the action will be designed to convey information in a credible way, to alter beliefs. The fact that customers will treat a firm that issues a better guarantee as if its product is better—and therefore be willing to pay a higher price—may affect the guarantee that the firm is willing to issue. Knowing that his selling his shares will convey a negative signal concerning his views of the future prospects of his firm, an entrepreneur may retain more of the shares of the firm; he will be less diversified than he otherwise would have been (and accordingly, he may act in a more risk averse manner).

A simple lesson emerges: some individuals wish to convey information; some individuals wish not to have information conveyed (either because such information might lead others to think less well of them, or because conveying information may interfere with their ability to appropriate rents). In either case, the fact that actions convey information leads people to alter their behavior, and changes how markets function. This is why information imperfections have such profound effects.

Once one recognizes that actions convey information, two results follow. First, in making decisions about what to do, individuals will not only think about what they like (as in traditional economics) but how it will affect others' beliefs about them. If I choose to go to school longer, it may lead others to believe that I am more able, and I will therefore decide to stay in school longer, not because I value what is being taught, but because I value how it changes others' beliefs concerning my ability. This means, of course, that we have to rethink completely firm and household decision making.

Secondly, we noted earlier that individuals have an incentive to "lie"—the less able to say that they are more able. Similarly, if it becomes recognized that those who walk up to the fifth floor to apply for

insurance are more healthy, then I might be willing to do so even if I am not so healthy, simply to fool the insurance company. If it becomes recognized that those who stay in school longer are more able, then I might be willing to do so, even if I am less able, simply to fool the employers. Recognizing this, one needs to look for ways by which information is conveyed *in equilibrium*. The critical insight as to how that could occur was provided in a paper with Michael Rothschild [1976]. If those who were more able, less risk prone, more credit worthy *acted* in some observable way (had different preferences) than those who were less able, less risk prone, less credit worthy, then it might be possible to design a set of *choices*, which would result in those with different characteristics in effect *identifying* themselves through their *self-selection*. One of the reasons that they might behave differently is that they *know* they are more able, less risk prone, more credit-worthy—that is there is asymmetric information. But it is only one of the bases for self-selection.

The particular mechanism which we explored in our insurance model illustrates how self-selection mechanisms work. People who know they are less likely to have an accident will be more willing to accept an insurance policy with a high deductible, so that an insurance company that offered two policies, one at a high premium and no deductible, one with a low premium and high deductible, would be able to sort out who were high risk and who low. It is an easy matter to construct choices which thus *separate*.

### *Monopoly and self-selection*

Analyzing the choices which arise in *full* equilibrium behavior turned out, however, to be a difficult task. The easiest situation to analyze was that of a monopolist.<sup>67</sup> He could construct a set of choices that would *differentiate* among different types of individuals, and analyzed whether it was profit maximizing for him to do so fully, or to (partially) “pool”—that is, offer a set of contracts such that several types might choose the same one. This work laid the foundations of a *general theory of price discrimination*. Under standard theories of monopoly, with *perfect information*, firms would have an incentive to price discriminate perfectly (extracting the full consumer surplus from each). If they did this, then monopoly would in fact be non-distortionary. Yet most models assumed no price discrim-

ination (that is, the monopolist offered the same price to all customers), without explaining why they did not do so, and argued that monopoly was distortionary. Our work showed how, given limited information, firms could price discriminate, but could do so only imperfectly. Subsequent work by a variety of authors (such as Salop [1977] and Adams and Yellen [1976]) explored a variety of ways by which a monopolist might find out relevant characteristics of his customers—the extent of discrimination limited by his ability to identify each person’s “surplus,” the maximum they would be willing to pay. (For an insurance company, the relevant characteristics are not only the likelihood of having an accident, but also the degree of risk averse, the premium that an individual would be willing to pay to divest himself of risk.) The economics of information thus provided the first coherent theory of monopoly.

### *Self-selection and competitive equilibrium*

The reason that analyzing monopoly was easy is that the monopolist could structure the entire choice set facing his customers. The hard question is to describe the full competitive equilibrium, that is a set of insurance contracts such that no one can offer an alternative set which would be profitable. Each firm could control the choices that he offered, but not the choices offered by others; and the decisions made by customers depended on the entire set of choices available. In our 1976 paper, Rothschild and I succeeded in analyzing this case.

Three striking results emerged from this analysis. The first I have already mentioned: under plausible conditions, given the natural definition of equilibrium, equilibrium might not exist. There were two possible forms of equilibria, pooling equilibria, in which the market is not able to distinguish among the types, and separating equilibria, in which it is. The different groups “separate out” by taking different actions. We showed that there never could be a pooling equilibrium—if there were a single contract that everyone bought, there was another contract that another firm could offer which would “break” the pooling equilibrium. On the other hand, there might not exist a separating equilibrium. The cost of separation was too great. Any putative separating equilibrium could be broken by a profitable pooling contract, a contract which

would be bought by both low risk and high risk types.

Second, even small amounts of imperfections of information can change the standard results, concerning the existence and characterization of equilibrium. Equilibrium, for instance, never exists when the two types are very near each other. As we have seen, the competitive equilibrium model is simply not robust.

Thirdly, and relatedly, we now can see how the fact that actions convey information affects the *equilibrium*. In particular, our analysis here reinforced the earlier analysis of adverse selection about markets not functioning well. In perfect information models, individuals would fully divest themselves of the risks which they face, and accordingly would act in a risk neutral manner. We explained why insurance markets would not work well—why most risk averse individuals would buy only partial insurance. There were numerous subsequent applications to other markets, reinforcing, for instance, the earlier conclusions concerning the limitations on equity markets. (The reason that the original owners of a firm might want to sell his shares was to “insure” himself against the risk of a bad outcome; an owner that believed that there was a smaller probability of a bad outcome would be willing to buy less insurance, i.e. to divest himself of fewer of his shares. Retention of shares can thus be thought of as a market sorting mechanism, the willingness to keep these shares a “signal” of the owners’ confidence.)<sup>68</sup>

The result was important not only for the insights it provided in the workings of an important set of markets in the economy, but because there are important elements of insurance in many transactions and the general principle that actions convey information, and that market transactions are greatly affected by this fact, has implications in a still wider variety of contexts. The relationship between the landlord and his tenant, or the employer and his employee, can be viewed as containing in it an insurance component; limitations on the ability to divest oneself of risk are important in explaining a host of contractual relationships.

### *Sorting, screening, and signaling*

In equilibrium, both buyers and sellers, employers and employees, insurance company and insured, lender and creditor are aware of the informational

consequences of their actions. Each side of the market needs to consider the consequences, e.g. of acting in a different way, or of confronting the other side of the market with different choices. In the case where say, the insurance company or employer or employee takes the initiative for sorting out applicants, self-selection is an alternative to examinations as a sorting device. In the case where the insured, or employee, or borrower, takes the initiative for identifying himself as a better risk, a better employee, a borrower more likely to repay, then we say he is *signaling*.<sup>69</sup> But of course, in equilibrium both sides are aware of the consequences of alternative actions, and the differences between signaling and self-selection screening models lie in the technicalities of game theory, and in particular whether the informed or uninformed (employee or employer, insured or insurance company) moves first.<sup>70</sup>

Still, some of the seeming differences between signaling and screening models arise because of a failure to specify a *full* equilibrium. We noted earlier that there were many separating contracts, but a unique separating equilibrium. We argued that if one considered any other separating set of contracts, then, say, in the insurance market, a firm could come in and offer an alternative set of contracts and make a profit; the original set of separating contracts could not have been an equilibrium. The same is true in, say, the education signaling model. There are many educational systems which “separate”—that is, the more able choose to go to school longer, and the wages at each level of education correspond to the productivity of those who go to school for that length of time. But all except one are not in *full equilibrium*. Assume, for instance, there were two types of individuals, a low ability and a high ability. Then if the low ability goes to 12 years of schooling, then any education system in which the high ability went *sufficiently* long—say more than 14 years—might separate. But the low ability would recognize that if it went to school for 11 years, it would still be treated as low ability. The unique equilibrium level of education for the low ability is that which maximizes his net income (taking into account the productivity gains and costs of education); and the unique equilibrium level of education for the high ability is the lowest level of education such that, if his wage corresponds to his productivity at that level of education, the low ability will still prefer to remain at his low



level of education rather than pretend to be more able by staying in school longer.<sup>71</sup>

The education system, of course, was particularly infelicitous for studying *market* equilibrium. The structure of the education system is largely a matter of public choice, not of market processes. Different countries have chosen markedly different systems. The minimum level of education is typically not a matter of choice, but set by the government. Within educational systems, examinations play as important a role as self-selection or signaling, though *given* a certain standard of testing, there is a process of self-selection involved in deciding whether to stay in school, to try to pass the examination.<sup>72</sup> For the same reason, the problems of existence which arise in the insurance market are not relevant in the education market—the “competitive” supply side of the market is simply absent. But when the signaling concepts are translated into contexts in which there is a robust competitive market, the problems of existence cannot be so easily ignored.<sup>73</sup>

### *Existence of equilibrium*

What are we to make of the problem of existence? Clearly, insurance markets *exist*, even if they are far from complete. To some extent, the market does exhibit instability. Rates vacillate enormously; as rates sometimes skyrocket and coverage is curtailed, the public clamors for reforms. Such periods are often followed by periods of relative stability, to be followed by another “crisis” in the market. Most states regulate rate setting, though at least partly for prudential reasons and this may help stabilize the market.

Moreover, though there is considerable evidence for the kinds of selection processes discussed above, there is also considerable evidence that the market is far from as rational as the theory would suggest. Most health insurance policies do not base premia on the number of children, though that is an easily observable variable which clearly affects the risk exposure. Many insurance companies do not use past experience as heavily as one would have expected in setting premia, i.e. there is less experience rating.

My own suspicion, however, is that the major limitation of Rothschild-Stiglitz is its assumption of perfect competition; competition is far more limited than we postulated; there are, for instance, significant search costs, and considerable uncertainty

about how easy it is to get the insurance firm to pay on a claim. Self-selection is still relevant, but the model of monopoly, or some version of monopolistic competition, may be more relevant than the model of perfect competition.

## THEORY OF CONTRACTS AND INCENTIVES

The work with Rothschild was related to the earlier work that I had done on incentives (sharecropping), besides the obvious way that both were concerned with problems of limited information, one focusing on selection effects and one on incentives. Both entailed equilibrium in “contracts.” The contracts that had characterized economic relations in the standard competitive model were extraordinarily simple: I will pay you a certain amount if you do such and such. If you did not perform as promised, the pay was not given. But with perfect information, individuals simply didn’t sign contracts that they did not intend to fulfill. Insurance contracts were similarly simple: a payment occurred if and only if particular specified events occurred.

The work on sharecropping and on equilibrium with competitive insurance markets showed that with imperfect information, a far richer set of contracts would be employed,<sup>74</sup> and thus began a large literature on the theory of contracting.<sup>75</sup>

In the simple sharecropping contracts of Stiglitz [1974b], the contracts involved shares, fixed payments, and plot sizes, and<sup>76</sup> more generally, optimal payment structures related payments to *observables*, inputs, processes, outputs.<sup>77,78</sup> Because what went on in one market affect others, the credit, labor, and land markets were *interlinked*; one could not decentralize in the way hypothesized by the standard perfect information model.<sup>79</sup> The theory thus served as the basis of the *rural organization* in developing countries.<sup>80</sup>

The basic principles were subsequently applied in a variety of other market contexts. The most obvious was the design of labor contracts.<sup>81,82</sup>

Payments, too, can depend on *relative* performance; relative performance may convey more relevant information than absolute performance. If a particular company’s stock goes up when all other companies’ stock goes up, it may say very little about the performance of the manager. In Nalebuff and Stiglitz [1983a, 1983b] we analyzed the design

of these relative performance compensation schemes (contests). One of the strong arguments for competitive, decentralized structures is that they provide information on the basis of which one can design better incentive pay structures than those which rely on the performance of a single individual only.

Credit markets too are characterized by complicated contracts. Lenders would specify not only an interest rate, but also impose other conditions (collateral requirements, equity requirements) which would have both incentive and selection effects.<sup>83</sup> Indeed, the simultaneous presence of both selection and incentive effects was important: in the absence of the former, it might be possible to increase the collateral requirement *and* raise interest rates, still ensuring that the borrower undertook the safe project.<sup>84</sup>

### *Incentives in market equilibrium*

Incentives are based on rewards and punishments. In modern economies, the most severe punishment that one can impose is to fire an individual.<sup>85</sup> But if the individual could get a job just like his current one, then there would be no cost. Good behavior is driven by earning a surplus over what one could get elsewhere. Thus, in labor markets, the wage must be higher than what the worker could get elsewhere (which may be zero, if there is unemployment); in the goods market, firms must feel a loss when they lose a customer because of a shoddy product, so the price must exceed the marginal cost of production. Thus, the long standing presumption that in competitive equilibrium price equals marginal cost *cannot* be true in markets with imperfect information. (See Shapiro and Stiglitz [1984], Shapiro [1983] and Klein and Leffler [1981].)

## **EQUILIBRIUM WAGE AND PRICE DISTRIBUTIONS**

One of the most obvious differences between the predictions of the model with perfect information and what we see in every day life is the conclusion that the same good sell for the same price everywhere. We all spend a considerable amount of time shopping for good bargains. The differences in prices represent more than just differences in quali-

ty (service). There are *real* price differences. Since Stigler's classic paper [1961], there has been a large literature exploring optimal search behavior. Stigler, and most of the search literature, took, however, the price or wage distribution as given. They did not ask how did it arise. Given the search costs, could it be sustained? For instance, if search costs are relatively low, one might have thought (if one bought the older theories) that markets would look very much like they would with zero search costs, in which case there would be no price or wage distribution. It is not surprising that given that information is costly, if there are shocks to the economy—the demand for a good goes up in some locale, so price there rises—that prices are not fully arbitrated instantaneously. But much of the wage and price dispersion cannot be related to such “shocks.”

Our analysis of efficiency wage theory provided an alternative explanation. We showed that it paid firms to pay more than they had to, e.g. to reduce labor turnover costs. But it might pay some firms to pay higher wages than others.

As I began to analyze these models, an important insight occurred: there could be a wage distribution even if all firms were identical, e.g. faced the same search costs. It was clear that even small search costs could make a large difference to the behavior of product and labor markets. This was a point that Diamond [1971] had independently made in a highly influential paper, which serves to illustrate powerfully the lack of robustness of the competitive equilibrium theory. Assume, as in the standard theory, all firms were charging the competitive price, but there were an epsilon cost of searching, of going to another store. Then any firm which charged epsilon/2 greater would lose no customers. It would thus pay him to increase his price. And it would similarly pay all other firms to increase their prices. But at the higher price, it would again pay each to increase his price. And price increases until the price charged is the monopoly price. Even small search costs thus lead even a market with many firms to charge monopoly prices. Work with Salop (Salop and Stiglitz [1977, 1982, 1987], Stiglitz [1979b, 1989c]), showed that in situations where there were even small search costs, markets would be characterized by a price distribution. If everyone were charging the same price, it would pay some firm either to raise his price, to exploit the high search costs customers who he would not lose, or to lower his price, to steal customers away from his

rivals. The standard wisdom that said that not everyone had to be informed to ensure that the market acted in a perfectly competitive manner was simply not in general true.<sup>86</sup>

## EFFICIENCY OF THE MARKET EQUILIBRIUM AND THE ROLE OF THE STATE

Perhaps the most important single idea in economics is that competitive economies lead, as if by an invisible hand, to a (Pareto) efficient allocation of resources, and that every Pareto efficient resource allocation can be achieved through a competitive mechanism, provided only that the appropriate lump sum redistributions are undertaken. It is these fundamental theorems of welfare economics which provide both the rationale for the reliance on free markets, and the belief that issues of distribution can be separated from issues of efficiency, allowing the economist the freedom to push for reforms which increase efficiency, regardless of their seeming impact on distribution; if society does not like the distributional consequences, it should simply redistribute income.

The economics of information showed that neither of these results was, in general, true. To be sure, economists over the preceding three decades had identified important market failures—such as the externalities associated with pollution—which required government intervention.<sup>87</sup> But the scope for market failures was limited, and thus the arenas in which government intervention was required were limited.

Early work, already referred to, had laid the foundations for the idea that economies with information imperfections would not be Pareto efficient, *even taking into account the costs of obtaining information*. There were interventions in the market that could make all parties better off. We had shown, for instance, that incentives for the disclosure and acquisition of information were far from perfect; imperfect appropriability meant that there might be insufficient incentives, but the fact that much of the gains were “rents,” gains by some at the expense of others, suggested that there might be excessive expenditures on information. One of the arguments for unfettered capital markets was that there were strong incentives to gather information; if one discovered that some stock was more valu-

able than others thought, and if you bought it before they discovered the information, then you would make a capital gain. This price discovery function of capital markets was often advertised as one of its strengths. But the issue was, while the individual who discovered the information a nano-second before any one else might be better off, was society as a whole better off: if having the information a nano-second earlier did not lead to a change in real decisions (e.g. concerning investment), then it was largely redistributive, with the gains of those obtaining the information occurring at the expense of others. Another example illustrates what is at issue. Assume hundred dollar bills were to fall, one each at the left foot of each student in my class. They could wait to the end of the lecture, then pick up the money; but that is not a Nash equilibrium. If all students were to do that, it would pay any one to bend down and quickly scoop up what he could. Each realizing that immediately picks up the dollar bill at his foot. The equilibrium leaves each no better off than if he had waited—and there was a great social cost, the interruption of the lecture.<sup>88</sup>

There are potentially other inefficiencies associated with information acquisition. Information can have adverse effects on volatility.<sup>89</sup> And information can lead to the destruction of markets, in ways which lead to adverse effects on welfare. We described earlier how the existence of asymmetries of information can destroy markets. Individuals sometimes have incentives to obtain information (creating an asymmetry of information), which then leads to the destruction of insurance markets, and an overall lowering of welfare. Welfare might be increased if the acquisition of this kind of information could be proscribed. Recently, such issues have become sources of real policy concern, in the arena of genetic testing. Even when information is available, there are issues concerning its use, with the use of certain kinds of information having either a discriminatory intent or effect, in circumstances in which such direct discrimination itself would be prohibited.<sup>90</sup>

Moreover, asymmetries of information were shown to be related to absent or imperfect markets. They help explain why the market for lemons, or the credit or labor or equity markets worked imperfectly. The fact that markets with imperfect information worked differently—and less well—than markets with perfect information was not, by itself, a damning criticism of markets. After all, informa-

tion is costly, and taking into account the costs of information, markets might be fully efficient. Stigler had essentially argued for this perspective, but without proof. Our research showed that this assertion—or hope—was simply not correct. Earlier work had established that when markets are absent or imperfect, market equilibrium might be *constrained* Pareto inefficient, that is, taking into account the absence of the market, everyone could be made better off.<sup>91</sup> Moreover, since asymmetries of information give rise to market power, and *perfect* competition is required if markets are to be efficient, it is perhaps not surprising that markets with information asymmetries and other information imperfections are far from efficient.

But while it was thus not surprising that markets might not provide appropriate incentives for the acquisition and dissemination of information, the market failures associated with imperfect information are far more profound. The intuition can be seen most simply in the case of models with moral hazard. There, the premium charged is associated with the *average* risk, and, therefore, the average care taken by seemingly similar individuals. The moral hazard problem arises because the level of care cannot be observed. Each individual ignores the effect of his actions on the premium; but when they all take less care, the premium increases. The lack of care by each exerts a negative externality on others.

The essential insight of Greenwald and Stiglitz [1986]<sup>92</sup> was to recognize that such externality-like effects are pervasive whenever information is imperfect or markets incomplete—that is always—and as a result, markets are essentially never constrained Pareto efficient.<sup>93</sup> In short, market failures are pervasive.

There were two other implications. The first was the non-decentralizability of efficient market solutions. The notion that one could decentralize decision making to obtain (Pareto) efficient resource allocation is one of the fundamental ideas in economics. Greenwald and Stiglitz showed that that was not in general possible. Again, a simple example illustrates what is at issue. An insurance company cannot monitor the extent of smoking, which has an adverse effect on health. The government cannot monitor smoking any better than the insurance company, but it can impose taxes, not only on cigarettes, but also on other commodities which are complements to smoking (and subsidies on substi-

tutes which have less adverse effects.)<sup>94</sup> Earlier work with Braverman [1982] had shown the consequences of this non-decentralizability, the interlinkage of land, labor, and credit markets in agrarian markets of developing countries.

Markets are also interlinked over time. *Intertemporal linkages* impair the efficacy of competitive processes, as we have already noted. Standard theory stated that if an employer does not treat an employee well, he simply moves to another firm. But informational asymmetries impair labor mobility, partially locking the employee into his employer, or the borrower into his creditor.<sup>95</sup> While with perfect information and perfect markets, some of the consequences of this reduction in ex post competition could be corrected by the intensity of ex ante competition, there is little reason to believe that is in fact the case.<sup>96</sup>

One of the sources of the market failures is *agency problems*, such as those which arise when the owner of land is different from the person working the land. The extent of agency problems—and therefore of market failures—thus depends on the distribution of wealth, as we noted earlier in our discussion of sharecropping. It is simply not the case that one can separate out issues of equity and efficiency.<sup>97</sup>

Moreover, the notion that one could separate out issues of equity and efficiency also rested on the ability to engage in lump sum redistributions. But as Mirrlees [1971] had earlier pointed out, with imperfect information, this was not possible; all redistributive taxation was distortionary. But this had important implications for a wider range of policies beyond simply the design of tax structures. It meant that interventions in the market which changed the before tax distribution of income could be desirable, because they lessened the burden on redistributive taxation.<sup>98</sup> Again, the conclusion: the second welfare theorem, effectively asserting the ability to separate issues of distribution and efficiency, was not true.<sup>99</sup>

In effect, the Arrow Debreu model had identified the single set of assumptions under which markets were (Pareto) efficient. There had to be perfect information, or, more accurately, information (beliefs) could not be *endogenous*, they could not change either as a result of the actions of any individual or firm, including investments in information. But in an information economy, a model

which assumes that information is *fixed* seems increasingly irrelevant.

### *Dysfunctional institutions*

As the theoretical case that markets in which information was imperfect were not efficient became increasingly clear, several arguments were put forward against government intervention. One we have already dealt with: the government too faces informational imperfections. But our analysis had shown that the incentives and constraints facing government differed from those facing the private sector, so that even when government faced exactly the same *informational* constraints, welfare could be improved upon.<sup>100</sup>

There was another argument, which held up no better. The existence of market failures—absent or imperfect markets—does give rise to non-market institutions. The absence of death insurance gave rise to burial societies. Families provide insurance to their members against a host of risks for which they either cannot buy insurance, or for which the insurance premium is viewed to be too high. But in

what I call the *functionalist fallacy*, it is easy to go from the observation that an institution arises to fulfill a function to the conclusion that actually, *in equilibrium*, it serves that function. Those who succumbed to this fallacy seemed to argue that there was no need for government intervention because these nonmarket institutions would “solve” the market failure, or at least do as well as any government. Richard Arnott and I [1991a] showed that, to the contrary, non-market institutions could actually make matters worse. Insurance provided by the family could crowd out market insurance; insurance companies would recognize that the insured would take more risk because they had obtained insurance from others, and accordingly cut back on the amount of insurance that they offered. But since the non-market (family) institutions did a poor job of divesting risk, welfare was decreased.<sup>101</sup>

The Arnott-Stiglitz analysis reemphasized the basic point made at the end of the last subsection: it was only under very special circumstances that markets could be shown to be efficient. Why then should we expect an equilibrium involving non-market institutions and markets to be efficient?