

# The productivity of failures

How a rejected paper generated a flourishing research programme.

## Ernst Fehr

In the late 1980s, I wrote a paper about involuntary unemployment, which I hoped would answer questions economists had been debating for decades. The paper was rejected, but developed into basic research that encompassed several disciplines — biology, psychology and economics.

Involuntary unemployment is a state in which unemployed workers are willing to work for less than the going wage, but nevertheless cannot get a job.

Common sense dictates that in periods of economic downturn, some workers without a job are involuntarily unemployed. In economics, there is a long-standing question as to whether involuntary unemployment can also occur in a state of equilibrium, that is, when no endogenous economic forces push for a change in the situation. A more complex and politically contested issue is whether an equilibrium with involuntary unemployment can occur in a free-market economy that is unconstrained by union power and minimum-wage legislation. The great British economist John Maynard Keynes argued that this is possible, but economists have grappled with this question for decades — and still remain divided.

My paper in the 1980s was entitled ‘Fair wages and unemployment’, and I thought it provided a convincing answer to these problems. Its central theme was that employees do not work hard if they are paid an unfair wage. Therefore, if the going wage represents a fair standard for reference, workers hired below this wage will be willing to accept an initial low wage offer, but are later — when employed — less likely to work hard. Hence, employers shy away from cutting employees’ wages or from employing workers below the going wage.

This hypothesis, that fairness concerns affect the functioning of the labour market, is an old one, but mainstream economics has largely disregarded this idea. Just like other people, most economists appreciate being treated fairly in their personal interactions — and would be deeply unhappy if paid less than what they consider to be fair. But when it comes to the modelling of economic affairs, fairness is not a currency that counts. In addition (like any other science), economics is characterized by conventions, one of the strongest being that a good model should generate predictions that are based on the



assumption that all agents maximize well defined objective functions. In principle, it is not that important what economic agents maximize, as long as they maximize something. But in practice, economists still have strong conventions regarding what constitutes a proper objective function — fairness has certainly not been among the objectives the convention legitimizes. Thus, I was not too surprised when several leading economics journals rejected my paper, although it was built on a framework where all agents maximize ‘something’. Unfortunately, the ‘something’ also included fairness goals.

I remember that one of the referees argued that the distance between the assumptions and the conclusions was “too small” to make the paper interesting. This argument reveals a lot about the psychology behind judging a paper’s merits. A paper has value when it creates a surprise — either empirically or in terms of theoretical insights. Science is often the prisoner of the psychology of surprise, and for theorists, elegance or mathematical beauty is also important. Despite all the rejections, I remained convinced that fairness concerns are important. In the end, it is not important whether a theory is simple or complicated, elegant or clumsy, beautiful or ugly — such considerations are of a secondary nature, alien to true science. Usually they are just a poor substitute for a simple fact — the absence of convincing empirical evidence. Therefore, I searched for ways to capture the issue of fairness empirically. In retrospect, this search was the major turning point of my scientific career and completely reshaped my research programme. I metamorphosed from a labour-market theorist into an experimental economist.

In the late 1980s there was already a large amount of experimental literature on fairness in bilateral bargaining, which started

with the economist Werner Güth’s path-breaking ultimatum game experiments. However, I was interested in whether fairness motives affect competitive markets. Experimental economists had already conducted thousands of competitive market experiments and the conventional wisdom that fairness is irrelevant in markets prevailed. But typically these experiments differed from the environment in my rejected theory paper because effort was assumed to be fully determined in the contract.

Fortunately, I received a grant enabling me to employ two excellent research assistants — Georg Kirchsteiger and Arno Riedl. Together, we designed an experiment which exactly implemented the economic environment assumed in my theory paper — in particular, that of non-contractible effort. We almost could not believe the data from the pilot experiment — we saw what we had only hoped for in our wildest dreams: fairness concerns strongly inhibited competitive forces and kept wages substantially above the competitive level. Moreover, as our theory predicted, wage reductions induced the employees to reduce their effort, rendering wage cuts unprofitable for the employers. Now we had the surprise we needed for publication.

My later work, and that of many others, indicated that fairness concerns permeate many other aspects of economic and social life — they affect competition and cooperation between and within firms, they influence international negotiations, the provision of public goods, exploitation of common property resources and they are the basis of many political conflicts. In addition, fairness norms probably played a decisive role in the evolution of human sociality. What began as research on involuntary unemployment turned into basic research about the nature of human altruism (see *Nature* 415, 137–140; 2002, and *Nature* 425, 785–791; 2003) and led to the concept of strong reciprocity — a term invented by Samuel Bowles and Herbert Gintis to describe the widespread human propensity to reward helpers and to punish cheaters. Today, this concept inspires anthropologists, zoologists, evolutionary biologists and evolutionary psychologists alike. I wish them success, and a few productive failures. ■

Ernst Fehr is at the University of Zürich, Institute for Empirical Research in Economics, Blümlisalpstrasse 10, CH-8006 Zürich, Switzerland.