Evaluating Immigration Policy Potentials and Limitations

Michael Fertig

February 2002
Evaluating Immigration Policy
Potentials and Limitations

Michael Fertig
University of Heidelberg and IZA, Bonn

Discussion Paper No. 437
February 2002

IZA
P.O. Box 7240
D-53072 Bonn
Germany
Tel.: +49-228-3894-0
Fax: +49-228-3894-210
Email: iza@iza.org

This Discussion Paper is issued within the framework of IZA’s research area Evaluation of Labor Market Policies and Projects. Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent, nonprofit limited liability company (Gesellschaft mit beschränkter Haftung) supported by the Deutsche Post AG. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public. The current research program deals with (1) mobility and flexibility of labor, (2) internationalization of labor markets, (3) the welfare state and labor markets, (4) labor markets in transition countries, (5) the future of labor, (6) evaluation of labor market policies and projects and (7) general labor economics.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available on the IZA website (www.iza.org) or directly from the author.
ABSTRACT

Evaluating Immigration Policy Potentials and Limitations*

Based on the ideas developed in the literature on the evaluation of active labor market policy, this paper provides a framework for the evaluation of key elements of immigration policy. To this end, the fundamental ingredients of evaluating policy interventions are explained and the specific case of immigration policy is analyzed. It becomes transparent that the evaluation of the effect of immigration policy is a particularly complex task since it requires unusually strong assumptions to hold a priori. These assumptions and possible reasons for their failure are discussed in detail. It is clarified that any violation of these assumptions renders the interpretation of the policy effects invalid. Finally, these insights are utilized for a critical review of the received literature.

JEL Classification: J61, H11

Keywords: Immigration policy, evaluation, identification assumptions

Michael Fertig
Department of Economics (Econometrics)
University of Heidelberg
Grabengasse 14
69177 Heidelberg
Germany
Fax: +49 6221 54 3640
Email: fertig@uni-hd.de

* The author is grateful to Christoph M. Schmidt for very helpful comments.
1 Introduction

During the last 50 years international migration flows have changed in intensity and composition to an extent\(^1\) which was not observed before. Since the end of World War II Europe as a whole underwent a transition process to an *immigration* region. The early European post-war migration experience has been dominated by migration streams from Europe’s South to Western and Northern Europe, with a clear focus on labor migration. In the course of time, the ethnic composition of immigration to Europe has changed dramatically. Europe as a whole has become a net receiving region, and the geographic and cultural distances to the immigrants’ countries of origin have increased significantly. These developments coincide with changes in the demographic and political situation in Europe. As a consequence, immigration policies of the destination countries are reacting to the new challenges by reshaping existing regulations, with the current discussion in Germany as a prime example.

On August 03, 2001 the Federal Interior Minister Otto Schily proposed a bill that intends to give Germany its first regulated immigration system ever. The proposed bill is motivated by the insight that “Germany is an immigration country” (Otto Schily) and that the country has to engage itself in the international competition for high-skilled migrant workers to pursue its own economic interests. The main directions of the proposed bill point towards (i) an active regulation of immigration by combining the work and residence permits with a point system for the selection of immigrants; (ii) an improvement of the existing integration policy by an extended system of language and culture courses; and (iii) a tightening of the existing right to apply for asylum. Regarding the first direction, high-skilled workers can obtain permanent residence and work permits if they score high enough in a point system which favors young and educated individuals with a sound knowledge of German and/or a special relationship to Germany. Students and less-skilled workers can initially receive a temporary permit which can be made permanent if they score enough points after some years.

Prior to this development, in August 2000 Germany introduced the so-called “green card” regulation to recruit high-skilled IT-specialists on a *temporary* basis. These latest developments only reflect a persistent phenomenon in yet unprecedented intensity. Historically, Germany has been an immigration country since the 1960’s despite the official wording. Even after the recruitment stop in 1973 Germany experienced a steady inflow of migrants, accompanied however, by a substantial outflow over time as well. The historical experience with immigration to Germany is described in many papers, among others SCHMIDT (1994), SCHMIDT (1996), ZIMMERMANN (1994) and ZIMMERMANN (1995).

---

\(^1\)For an overview see CHISWICK AND HATTON (2001) for the case of Europe and BORJAS (1999) for the case of the US.
On the other hand, in the current negotiations regarding the enlargement of the European Union towards Central and Eastern Europe the extension of the free movement of labor agreement towards the prospective member states is heavily discussed. Many people in Germany, and also some economists (e.g. HÖNEKOPP AND WERNER (1999) or SNN (2000)), express serious concern about the possibility of a massive inflow of foreigners after the enlargement of the EU. Although the weight of the available evidence (see e.g. FERTIG (1999) and FERTIG AND SCHMIDT (2000a)) clearly suggests that this fear has no substantial grounds, it is very likely that the extension of the free movement agreement towards the new member states will be postponed for a transitory period of five to ten years.

In any event, there will likely be immigration to Germany in the future, and given the demographic and labor market developments in Germany there is also a dire need for further immigration. There are many possibilities for the regulation of these future immigration flows. Ample precedence for these possibilities is provided by the many different immigration policy regimes operating in different countries all over the world. This nexus is precisely the object of interest in this essay. In the context of this paper the label “immigration policy (regime)” comprises all policy measures aiming at the regulation of the entry process of immigrants as well as all attempts at their integration into the destination country’s society in the period directly after their arrival. This definition does, however, not include similar policy measures regarding asylum seekers. At the present time, there has been no conceptual attempt in the migration literature at formally evaluating different immigration policy regimes regarding their impact on observed migration flows to a specific country. By contrast, the received literature either takes it for granted that regulating entry is exerting a marked effect on immigration flows, or collects rather cursory evidence to this effect.

As with any other serious evaluation attempt, the impossibility of collecting experimental evidence implies that evaluating the effect of immigration policy requires strong assumptions to hold a priori. These assumptions are discussed in detail below. It becomes transparent that any violation of these assumptions renders the interpretation of the observed phenomena as causal effects of immigration policy invalid. This paper, therefore, provides a discussion of the problem of evaluating immigration policy resting on recent insights on causal analysis in econometrics and statistics. The concrete aim of the paper is to explain the fundamental issues of evaluating policy interventions and to analyze the specific problems of evaluating immigration policy interventions. For this purpose the principal strategy is to introduce an analogy to the evaluation of active labor market policy (ALMP). On the basis of this analogy it becomes transparent that the evaluation of the effects of immigration policy is a comparable, though more complex task than the evaluation of ALMP.

The first problem in this endeavor arises from the fact that the objec-
tives of immigration policy of a given country are often not completely clear. At least theoretically, ALMP is typically pursued to bring unemployed individuals back into work or to enhance the income situation of disadvantaged workers. Of course, some measures might implicitly also be pursued because they demonstrate activity and are regarded by politicians as a possibility to be relected. But it is certainly fair to gauge the success of ALMP mainly by their economic effects. By contrast, the aims of immigration policy are often of a variegated and certainly not of a exclusively economic nature. ROTTÉ (1998), for instance, provides a discussion of the variety of motives and objectives which might be detected behind immigration policy in Germany.

Typically, proponents of a specific immigration policy offer a set of non-economic arguments, like the achievement of cultural or religious homogeneity or diversity, respectively, or the avoidance of ethnic conflicts. Economic goals of immigration policy are e.g. fostering economic growth in the destination country, increasing the size or altering the composition of the population or the labor force, or enhancing the provision of the economy with human capital. Existing immigration policy regimes are motivated, explicitly or implicitly, by a mixture of such economic and non-economic objectives. However, distinctions like that are somewhat artificial since, for instance, the achievement of diversity is also a possible economic argument. LAZEAR (2000), for instance, argues that it is possible for an immigration country to realize a return from diversity in immigration flows since there may be a high reward on interactions between people with different backgrounds regarding creativity, information, cognitive ability or motivation. In the public discussion it is e.g. sometimes argued that the Asian immigrants in the US exhibit a different work ethic than natives or other immigrant groups which is perceived as one of the keys for their success.

There is a vast body of literature on immigration policy either for a specific country or on a synoptical basis for a set of countries. Section 2 provides a brief overview on this literature together with a discussion of the tasks of immigration policy. However, the majority of the papers within this literature is predominantly descriptive in nature and does not intend to follow a rigorous conceptual framework for an assessment of the effects of different policy regimes. Typically, the specific experience of any country is taken to be an episode too idiosyncratic to include it in a all-encompassing formal framework. From the perspective of economic policy, however, it is of prominent relevance to provide an answer to the question what would have happened to observed immigration – i.e. its magnitude and/or composition – to a specific country if the immigration policy regime of this country had indeed been different. This is the (implicit or explicit) counterfactual question of any empirical study on the effect of immigration policy, although no previous study explicitly discussed this aspect. The unobservability of this situation constitutes the fundamental evaluation problem, its solution must rest in a credible construction of such a comparison.
This paper, therefore, explains the principles of evaluating immigration policy in the light of the literature on the evaluation of public policy interventions in section 3. After a clarification of the relevant counterfactual question, the principal problems of defining an adequate outcome measure, choosing appropriate identification assumptions and measuring the "treatment effect" of different policy regimes will be discussed. Furthermore, this section turns the attention back to several important contributions in the empirical literature on immigration policy regimes. These papers are reconsidered again in the light of the discussion of section 3. Finally, section 4 offers some conclusions.

2 Immigration Policy – Literature and Analogy

This section provides an overview on existing immigration policy regulations currently in operation in the major immigrant receiving countries. Starting with a brief survey of the received literature on immigration policy the fundamental problem of evaluating immigration policy is then discussed, using an analogy to the evaluation of active labor market policy.

2.1 Survey of Literature

When providing a brief survey of the empirical literature on immigration policy it seems advisable to concentrate on a set of selected, particularly important contributions. The papers discussed in what follows are only a small subset of the vast body of contributions to this topic. However, taken together there are two strands in the received literature on immigration policy. The first strand analyzes data on a (semi-) aggregate level together with developments in the institutional settings of immigration policy over time. The second, and smaller, strand utilizes individual-level data to analyze the effects of a specific feature of a given immigration policy regime.

In the first group, most of the empirical papers present (semi-) aggregate evidence of the effects of immigration policy on a rather descriptive level. The papers\(^2\) analyze the skill-, country-of-origin- and age-composition of immigration flows and paint a variegated picture of immigration outcomes under different policy regimes. They do not, however, provide any quantitative evidence on the effect of immigration policy on a specific economic outcome measure. While these analyses are insightful and informative, it is obvious that one would benefit from the additional insight not aimed at: what would

the observed phenomenon looked like, if all countries under study pursued a specific immigration policy instead of the actual in operation?

In the second group, comparable to the rest of economic migration research, the analysis of the “classical” immigration countries, i.e. Australia, Canada and the US, has taken center stage. In the United States\(^3\) immigration is clearly dominated by kinship migration, i.e. induced by family reunification considerations. Skill-based immigration – that is, entry based on the provision of specific skills by migrants – is small compared to kinship migration, albeit high in absolute numbers compared to many European countries. On the contrary, the current immigration policies of Canada and Australia are dominated by selection mechanisms which reward formal skills rather than family ties. The differences in immigration policy between these countries as well as the different categories of migrants provide the basis for many of the analyses pursued in the second group of the empirical literature.

Borjas (1993) compares the educational attainment and the labor market performance of immigrants to the United States and Canada using the 1970 and 1980 census waves for each country. He reports that the average skill-level of different immigrant cohorts to Canada is higher than that of immigrants to the United States and attributes this finding to the more skill-based immigration policy regime operating in Canada. Furthermore, the estimation results of earnings regressions for both countries suggest that immigrants to the United States exhibit a higher earnings disadvantage upon entry relative to natives than immigrants to Canada. A decomposition analysis of the upon-entry earnings disadvantages of migrants demonstrates that the difference between migrants to the US and migrants to Canada can be explained by the differences in the national origin composition of immigration flows to both countries. Borjas concludes that the Canadian point system is not able to attract more skilled immigrants from a given country of origin.

Duleep and Regets (1996) aim at analyzing the effect of admission criteria on the labor market success of migrants in the US\(^4\). For this purpose they compare the relative earnings performance of kinship immigrant men to the relative earnings performance of skill-based immigrant men conditionally on observable factors like education and labor market experience in a regression framework. The authors conclude that although kinship migrants display a higher earnings disadvantage upon entry relative to natives than skill-based immigrants, this disadvantage vanishes over time. They suggest that this catching up process is related to a higher investment into human capital by kinship migrants since they display a lower degree of transferability of their country of origin specific human capital.

For the case of Australia, Chiswick and Miller (1992) estimate the determinants of unemployment conditional on immigrant group and other ex-

\(^3\)Borjas (1999) provides an overview on the US literature.

\(^4\)A similar analysis is conducted by Jasso and Rosenzweig (1995).
planatory factors. Estimation results suggest that there is no statistically significant difference in the unemployment situation between immigrant groups once one controls for education and other covariates.

For the case of Canada, Wright and Maxim (1993) provide an analysis of immigrant earnings conditional on immigrant status and human capital variables. The authors introduce an immigrant “quality” measure in their analysis by estimating the upon-entry earnings differential of different immigrant groups compared to native Canadians. The authors conclude that the skill-based selection system works better than other systems (e.g. family reunification) if success is measured by the upon-entry earnings differential. The empirical approaches and results of the contributions by Borjas (1993), Duleep and Regets (1996), Chiswick and Miller (1992) and Wright and Maxim (1993) will be reconsidered in chapter 3.

A different aspect of immigration policy is analyzed by Bauer, Lofstrom and Zimmermann (2000). The authors, using the 1995 wave of the ISSP, provide a cross country comparison on the perception of immigrants in different European and Non-European countries with a special focus on the influence of immigration policy on the attitudes towards minorities by the native population. The authors conclude that in countries with a more skill-based immigration policy, like Canada, natives have a more positive attitude towards immigrants than in countries with other policy regimes.

In the received literature, the problem of illegal migration is analyzed mainly from a theoretical perspective (see e.g. Todaro and Maruszko (1987), the special issue of the Journal of Population Economics (1999) or for a more recent contribution Entorf (2000)). The amount of empirical evidence on this topic is rather scarce. One exception is the paper by Cobb-Clark et al. (1995) analyzing the effect of the Immigration Reform and Control Act of 1986 in the United States on the wages of immigrant workers in the manufacturing sector. This act imposed sanctions on employers who hire illegal immigrants and legalized many long time illegal immigrants in the US. The evidence presented suggests that there is a small negative effect of employer sanctions and small positive effect of legalization on workers’ wages.

Many papers on immigrant performance demonstrate the high relevance of language skills for the success of immigrants in the destination country’s labor market (see e.g. Chiswick (1998) and Chiswick and Miller (1999)). It is widely agreed that language skills are an important aspect of integration policy. Cobb-Clark et al. (2001) explore the role of post-migration investment in formal and informal human capital by immigrants for the case of Australia. The empirical results presented by the authors suggest that these investments play a substantial role for the labor market performance of migrants. Consequently, public assistance for the acquisition of job search skills and formal education seem to contribute to the success of immigrants in Australia.
This brief overview demonstrates that the received literature of economic migration research addresses a variety of aspects regarding immigration policy. These contributions combined with the research conducted on the impact of immigration on the receiving countries' economy deliver useful insights for an adequate regulation of future immigration in the destination countries. However, it is very difficult to pin down stable relationships between specific policy regulations and measurable outcomes reflecting the objectives of immigration policy. In this paper it is argued that the reason for this difficulty is the missing common frame of reference for the different studies conducted. The following sections, therefore, outline such a framework by discussing the different elements necessary for any serious evaluation attempt. To this end it is regarded as helpful to clarify the fundamental challenges of evaluating immigration policy. This is done by analogy to the evaluation of ALMP.

2.2 The Analogy

As a consequence of an exploding literature on the subject during recent years, the literature on the evaluation of active labor market policy is rather mature. If not for each and every application, at least at the conceptual level the potential and limitations of attempts to evaluate such measures are understood quite well. This is not the case for immigration policy, though. Therefore, this literature can serve as a clarifying background to compare the similarities and differences in the evaluation of both types of policies.

Active Labor Market Policy

Consider the case of non-experimentally evaluating the effect of a specific measure of active labor market policy, e.g. a training measure or a wage subsidy for unemployed workers. Suppose there is a pool of individuals willing to participate in a specific program. For an economist it is natural to think that the individuals who apply to participate in the measure have based their decision on a thorough cost-benefit-comparison of the program. If they apply to the program they, therefore, must expect a positive return from participation. Otherwise, they would refrain from an application and search for a new job on their own. Consequently, the observable as well as unobservable characteristics of applicants to the program are hardly a random sample of the population. However, there might also be a (presumably smaller) amount of individuals who do not apply voluntarily but are forced to do so. This might be the case, if they e.g. would otherwise run the risk of loosing their benefit payments.

In a second step, from this pool of potential participants the labor offices typically select those individuals who will be granted admission to this measure according to the overall guidelines set out by the legislator. If more than the (more or less fixed) maximum amount of people for the measure fulfill the formal requirements for participation, then the responsible persons
at labor offices have to decide which of the eligible unemployed should be granted admission to the program and which not. For the observer this selection process is usually a black box. The details of the decision process in the labor offices as well as their internal guidelines upon which potential participants should be chosen for a measure are typically unobservable and to the analyst must remain highly speculative. It seems quite natural to suppose that the employees of the labor offices base their decision on a personal assessment of the ability of potential participants to be successful in a specific program. Therefore, there is a high probability that the labor offices choose those candidates for the measure who seem to display the highest motivation or cognitive abilities. However, it is far from being guaranteed that the selection process is operating in such a way since the details of this decision process are usually unobservable.

The result of this black box, however, is observable. After the selection process has been completed there is a group of unemployed workers participating in the program (the so-called treatment group) and a group of non-participants searching for a job without the support of the measure. The latter individuals constitute the pool of a potential comparison group. To gauge the effect of the intervention on the participants, one sensibly concentrates on the labor market success of both groups a sufficiently long time period after the program is completed. The difference in the labor market success between both groups (treatment and comparison group) measured by the value of a suitably defined outcome measure can than be used to judge the effectiveness of the program. To construct a credible comparison, though, by some strategy the researcher has to select an appropriate comparison group from the reservoir of potential comparisons. This, however, requires that convincing identification assumptions – stating clearly what is the appropriate comparison group – are invoked and that a suitable treatment parameter is defined.

These identification assumptions are necessary to construct an observable counterpart for the unobservable counterfactual situation. In the case of ALMP the counterfactual situation is implied by the question: What would have happened to the labor market success of participants if they had not participated in the measure? In this case, convincing identification assumptions must be able to “correct” for the presumably positive (double) selection of participants, i.e. the fact that a (presumably) positive self-selection in the application and an also (presumably) positive selection in the labor offices results in a non-random participant group. This is a necessary prerequisite to facilitate the attribution of a causal effect of the program on the chosen outcome measure in an observational or non-experimental study. However, if the latter prerequisite does not hold then it is only possible to attribute a causal effect of both the program and the specific selection processes at work on the outcome measure. This is exactly what one would like to avoid.

*Immigration Policy*
Now consider the case of evaluating immigration policy measures. There are strikingly large similarities but also considerable differences generating additional problems. Suppose there is a pool of individuals willing to immigrate into a given country. Again it seems natural to think that these individuals have based their decision on a thorough cost-benefit-comparison. They compare the net present value of the expected utility streams from migrating to a specific destination country $i$ with the net present values from migrating to all other destination countries and the net present value of staying at the origin. Consequently, for people applying for admission to country $i$ their net present value of immigrating to this country must be the highest. In other words, they must expect a positive return from immigrating to a specific country. If this were not the case, they would refrain from doing so and stay in the country of origin or go elsewhere. As a result the observable as well as unobservable characteristics of those people who eventually decide to emigrate from their origin country are also hardly a random sample of the population of the relevant country.

As it is the case for applicants to a labor market program, it is conceivable that the immigration policy of a specific country itself might have an effect on the “supply” of potential migrants, i.e. the pool of individuals willing to immigrate into the country. This might be the case, if the policy is able to serve as a signal for migrants that the specific selection mechanism, e.g. one which is awarding specific skills, is equivalent to a high return for those skills on the destination country’s labor market. If this supply-side effect of immigration policy is negligible, though, the magnitude and the composition of the pool of potential migrants can be taken as exogenous to immigration policy. However, there is clearly also an amount of individuals who do not emigrate voluntarily from their origin country but are forced to do so due to civil wars or famines. These are refugee migrants for which the freedom of choice is definitely limited and the decision process will certainly not adhere to individual utility maximization. Perhaps, they are at least able to decide to which country they emigrate, but this is far from being guaranteed. The latter people, however, usually apply for access to a country via humanitarian channels and the discretion of immigration offices is limited by international regulations like the Geneva convention on the status of refugees.

In a second step, from the pool of potential immigrants (those not applying admission via humanitarian channels), the immigration offices typically select those individuals who will be granted admission to the country according to overall guidelines set out by the legislator. If more than the (more or less fixed) maximum amount of people for immigration fulfill the formal requirements to access the country, then the responsible persons at the immigration offices have to decide which of the eligible individuals should be granted admission and which not. Again, this selection process is usually a black box for the observer. The details of the decision process in the immigration offices as well as their internal guidelines upon which potential immigrants should be chosen for admission are unobservable to the analyst.
in practice and must remain highly speculative. As it is the case for ALMP, it seems quite plausible to suppose that the employees of the immigration offices base their decision on a personal assessment of the ability of potential immigrants to be successful in the destination country. Therefore, there is a high probability that the immigration offices choose those candidates for admission who seem to display the highest motivation or cognitive abilities. However, it is far from being guaranteed that the selection process is operating in such a way.

The result of this black box, however, is again observable. After the selection process has been completed there is a set of individuals for which admission to the country has been granted (denote them again as the treatment group) and a set of individuals which has not been admitted. The first group, the new immigrant cohort in the destination country, comprises individuals from different countries of origin with different individual backgrounds regarding the details of the immigration motives, education, labor market experience, knowledge of the destination country’s language, motivation etc. This group might then be prepared for their access on the destination country’s labor market by integration measures like language courses or programs which impart job search skills. Once they have entered the labor market of the destination country one is usually able to observe the success of this treatment group in terms of a suitably defined outcome measure. Yet, this is the point where the analogy to ALMP ends and the additional differences of evaluating the effects of immigration policy come to bear.

Contrary to ALMP, the “treatment” of immigration policy is either the selection process in the immigration offices itself or the combination of this selection process with the upon entry integration measures. For an assessment of this treatment it is of course necessary to define a suitable treatment parameter as well. Furthermore, it is decisive to find a convincing comparison group. This is a complex task because one could not observe the non-migrants, i.e. those individuals who initially applied for admission to the country but were not accepted. Defining such a comparison group is equivalent to invoking convincing identification assumptions as it was the case with evaluating ALMP.

However, contrary to the evaluation of ALMP these assumptions should not simply correct for the selection of immigrants, since this selection is (part of) the phenomenon one is interested in. Rather, convincing identification assumptions must be able to reveal the results of the second selection process (i.e. that of the policy) net of the effect of the first self-selection process (i.e. that of the migrants themselves). The attribution of a causal effect of a specific immigration policy on the success of immigrants to a specific country is possible, if and only if this task is solved convincingly.

To organize ideas, consider as a completely hypothetical benchmark situation the case of a lottery, i.e. a distribution of the (more or less) fixed
amount of work permissions among all individuals willing to immigrate by chance. Given that operating an elaborate selection system is more costly than a lottery, it is a necessary condition for a specific immigration policy to outperform at least the lottery system in order to have a chance of being efficient. This means, that as a minimum prerequisite, any real world immigration policy regime should be able to select more successful migrants than the lottery would. Since no country is operating a lottery system such a comparison situation is not observable. Nevertheless, this benchmark provides the conceptually ideal “no active immigration policy” regime from which all actual regimes distinguish themselves.

The economic success of a migrant cohort in the destination country is the result of the interplay between observable and unobservable characteristics (such as the details of the motivation to immigrate, cognitive ability, the degree of transferability of origin country-specific skills, motivation, return intentions etc.) upon which admission was granted as well as the economic and institutional environment on the destination country’s labor market. The admission process itself upon which immigrants entered the country may hardly impinge upon their economic success separately. Rather, it exhibits an indirect influence via the selection criteria. In the light of this quite intuitive discussion the next section provides a more formal discussion on the necessary elements of any serious evaluation study and suggests the construction of a comparison group which – under specific assumptions – is able to provide a solution to the fundamental evaluation problem.

3 The Principles of Evaluating Immigration Policies

Every empirical study is confronted by a counterfactual question\(^5\). In the case at hand the counterfactual question of an evaluation study of immigration policy is how the immigration experience to a given country – measured by an adequate outcome measure – would have been, if the immigration policy regime of this country had been different. The fundamental problem is that this counterfactual situation is not observable, since one observes a given country at a given point in time only once. This means that only one particular policy, and one particular immigration cohort with one particular composition are observed together.

It is, therefore, necessary and the decisive point for any evaluation study to invoke identification assumptions in order to construct an observable coun-

\(^5\)For a survey see Heckman et al. (1999) or Blundell and Dias (2000). Schmidt (1999) provides an in-depth discussion of the problem of constructing a reliable counterfactual situation for the case of evaluating active labor market policy interventions on the individual level. Fertig and Schmidt (2000b) discuss the principles of evaluating labor market policy on a semi-aggregate level.
terpart of this unobservable situation. As a preceding step it is necessary to define an adequate outcome measure, a task which is particularly problematic in the case of evaluating immigration policy. This task will be tackled in the next subsection.

3.1 The Outcome Measure

The first step in any serious evaluation attempt is to choose an appropriate outcome measure. In this context it is also necessary to distinguish between effectiveness and efficiency of a policy measure. A specific measure is deemed effective, if the aims of the policy intervention are achieved at all and it is called efficient, if those aims are achieved by the smallest conceivable efforts. In particular, there must be no feasible alternative achieving the same aim at lower cost. For the evaluation of the efficiency of a policy measures it is also necessary to take into account possible unintended or adverse side effects.

For example, in the case at hand the substitution of low-skilled native workers by immigrants or long-term strains of the public health or pension systems might be unintended adverse side effects. Often the aims of immigration policy are mainly of economic nature, for instance fostering economic growth in the destination country by attracting otherwise unavailable unskilled labor. Then, it might also be possible to achieve this goal by lowering barriers to trade with countries which produce goods and services containing mainly low-skilled labor. Evaluating efficiency is a notoriously difficult task since it is nearly impossible to determine all relevant cost and it is, therefore, usually neglected in empirical studies. Most commonly the focus of empirical studies lies on the effectiveness of policy interventions.

Choosing an adequate outcome measure and measuring the cost of a specific immigration policy regime is closely related to the economic impact of immigration. This impact unfolds in an indirect fashion via market reactions and its measurement is therefore a complex task. Additional immigration shifts the relevant labor supply curve outward. The direct consequences, in terms of employment and wages for the relevant groups of workers are a matter of the relative own elasticities of demand and supply and of the set of elasticities of complementarity with all other production factors.

However, the additional labor supply effect is only one side of the medal, since product demand, and thus labor demand (on all other sub-markets) might be affected positively. On balance, it might not be the case at all that immigration harms any group of native workers via the crowding out that the constant-output reasoning typically applied seems to suggest. The common problem of empirical (i.e. non-experimental) research on this topic is the fact that additional immigration does not vary randomly across time and space but is rather the outcome of systematic forces. Thus, comparing the relevant economic outcome measures across regions may confuse the im-
pact of immigration with the underlying reason making the area particularly attractive.

Given the difficulties in measuring cost adequately, this paper does not attempt at evaluating the efficiency of immigration policy as well. Rather, the focus of the succeeding analysis lies completely on the evaluation of effectiveness. However, there is no guarantee that an effective immigration policy is also efficient. Moreover, effectiveness of immigration policy regimes is analyzed solely from the perspective of the destination country neglecting negative effects for the origin country (e.g. the loss of high-skilled labor for the origin country, the so-called “brain drain”). Effectiveness is, therefore, measured as the degree of reaching the (implicitly or explicitly) declared aims of the immigration policy of a destination country.

In the case at hand a natural candidate for an adequate outcome measure is the “success” of immigrants entering the country under a specific immigration policy regime. Success in terms of economic objectives, however, can be measured in different ways. In the long run economic success means that there is a welfare gain for the destination country economy. Welfare gains may be approximated (if only imperfectly) by significantly higher growth rates (per capita) due to a specific immigration policy regime. Given the data situation this approach does not seem feasible since this would require data over a sufficiently long time horizon, say 20-30 years.

In contrast to such a long term concept, a short run approach in assessing the ability of immigration policy to bring forth successful migrants would be to look at the average skills of a cohort of migrants. Since the pool of high skilled labor is commonly acknowledged as one of the major determinants of future economic growth (see e.g. Borjas (1999)) it seems natural to evaluate immigration policy by comparing the skills of immigration cohorts under different policy regimes. However, this approach suffers from the problem that human capital acquired in a specific origin country is usually not fully transferable to the destination country’s labor market. Moreover, initial differences in observable skills may not matter very much for economic performance and migrants’ contribution to economic growth (e.g. since it might be the unobservable traits common to all immigrants – motivation and perseverance – which matter). Finally, migrants might typically close a large initial gap faster than a small disadvantage, since investment into country-specific skills is less costly in terms of forgone earnings (as indicated by the results of Duleep and Regets (1996).

As the solution the middle way seems to be appropriate. A medium term concept in this endeavor would be based on the argument that migrants who are employed with relatively high earnings and, therefore, perform well on the destination country’s labor market are determinants of long run economic growth as well. Moreover, selecting migrants with a relatively high labor market performance is closely connected with selecting high skilled migrants but
also means that the skills of these migrants must be widely transferable to the destination country’s labor market.

It is, therefore, argued to assess the effectiveness of immigration policy regimes by using the labor market performance of immigrants in terms of wage or employment aspects under different regimes as an outcome measure. With this outcome measure the counterfactual question stated above could be put in the more precise form: How would have been the labor market performance of a given cohort of immigrants in a given country if the immigration policy regime had been different? Since this counterfactual situation is not observable one has to invoke adequate identification assumptions to construct an observable counterpart. The following section, therefore, discusses possible assumptions suitable for identification purposes.

### 3.2 Identification Assumptions

Valid identification assumptions are assumptions that, in principle, allow the estimation of the parameters of interest with growing precision if the sample size increases. Clearly, since it is not possible to observe the performance of a given cohort of immigrants at a given point in time under different policy regimes, collecting more or even better data would never suffice to identify any parameter of interest. Instead, one has to invoke an assumption which yields a comparison between immigrant cohorts under different policy regimes where both situations ideally differ in only a single aspect, the specific policy regime. Such assumptions are not testable and, consequently, have to be assumed to hold a priori.

In the evaluation literature, e.g. regarding active labor market policy, social experiments are usually recognized as a convincing identification strategy (see Heckman (1996) and Schmidt (2000)). The decisive feature of a social experiment is the randomized assignment of individuals who are willing to participate in a specific measure into a treatment and a control group. However, in the case of evaluating immigration policy considering such a randomized assignment will not be feasible, since it is the explicit aim of all immigration policy regimes to select those individuals which will be the most suitable to achieve the aims of the policy, without any room for experimentation. Therefore, with this key feature of social experiments being not feasible, the whole approach of an experiment is not suitable to solve the evaluation problem.

In terms of a formal analysis, however, the concept of a randomization mechanism provides a sensible theoretical benchmark for the assessment of the effectiveness of immigration policy. The fiction of a randomized assignment of potential migrants to a set of potential destination countries may serve as a sensible reference situation to which the success of actual policy regimes might be compared. Specifically, as outlined above, a specific immi-
igration policy might be termed successful if it is able to select immigrants out of the pool of potential migrants which are more successful in terms of their relative labor market performance than what would have been the result of a randomized immigration mechanism, for instance by a lottery of work permissions for different countries. Since such a reference situation is not observable one has to concentrate on direct comparison between different policy regimes, rather than on assessments of each regime in contrast to the benchmark situation⁶. Again, the conceptual requirement is that one has to construct a comparison situation by invoking identification assumptions.

In general, the central idea of “comparing the comparable” laid down in the evaluation literature in this context means that one wishes to compare the effect of different policy regimes conditional on a given supply of potential migrants. Therefore, the decisive point for any identification assumption to be convincing, is the ability to disentangle the double-selection processes in immigration. This means, that a convincing identification strategy must be able to discriminate between the effect of the self-selection of migrants (the supply, for short) and the effect of the selection by the immigration policy regime. This provides us with the possibility to assess the “quality” of a specific identification strategy by clarifying which assumptions have to hold in order to to assign a causal effect to immigration policy and to assess whether these assumptions are plausible in terms of economic considerations.

One possible and easily implementable approach would be a comparison across time (before-after-comparison) for a given country. This means that the performance of immigrants before and after a specific intervention in immigration policy, e.g. a change of policy from family reunification towards a skill-based selection, is compared and any change in this performance is causally attributed to the policy change. This requires the assumption that the performance of immigrants to that country, conditionally on a set of observable factors, would have been unchanged if the immigration policy intervention had not occurred. This is certainly a very restrictive assumption since there is typically a host of possible other factors changing with the regime switch, e.g. a changing demand for labor. Furthermore, any upturn in the business cycle, for example, that affected the labor market performance of immigrants in the specific country positively, would be attributed erroneously to the policy change yielding an overstated effect of the policy intervention. This cyclical sensitivity of a before-after-comparison is certainly one of the most convincing argument against proceeding in such a way. Finally, disentangling the double-selection process with this approach requires the assumption that there is no difference in the supply of migrants over time. This again seems to be a very restrictive assumption since the country-of-origin composition of migrants worldwide is changing over time and the observable as well as unobservable characteristics of these migrants are in all likelihood changing too.

⁶See Klüve (2001) for an elaborate discussion of the necessity and importance of choosing the correct comparison situation in any evaluation attempt.
Another possible identification assumption would invoke a comparison across space (cross-section-comparison), i.e. the comparison of a group of migrants conditional on a set of observable characteristics between different countries with different policy regimes at a given point in time. This requires the identification assumption that the relative labor market position of immigrants in country \( j \) had been equal to the relative labor market position of immigrants to country \( i \) \((i \neq j)\) if the policy regime operating in country \( j \) had been equal to the policy regime operating in country \( i \) and that there is no difference in the supply of migrants to the different countries. Again, this is not a very convincing assumption since the pool of potential migrants to a given country is hardly a random sample of the population of potential migrants.

Furthermore, this identification assumption is vulnerable regarding the business cycle, too. However, in this case the problem is the position of the relevant countries in the business cycle. Finally, this identification assumption requires that the labor market performance of the native comparison groups in all countries under consideration is equal, since a given group of migrants, with a specific set of observable and unobservable characteristics, will be performing relatively better in a country with an, on average, less successful native comparison group.

This paper, therefore, suggests as a solution to these problems a combination of the afore mentioned identification assumptions. This combination takes the form of a difference-in-differences comparison, i.e. a comparison across time and space. Suppose we aim at assessing the effect of a change in immigration policy in a specific country on the value of a specific outcome measure, i.e. we wish to evaluate effectiveness of policy change by a mean effect of the policy change on those selected by the new policy. Specifically, consider the case of two countries (\( A \) and \( B \)) with a comparable immigration policy regime before a regime change (e.g. currently Germany and Israel where admission depends on descent or Canada and the US until the 1980's with the focus on family reunification). Furthermore, assume that the relative earnings performance of the migrant cohorts is an adequate outcome measure.

Throughout this analysis there are two maintained assumptions: (i) the respective policy in both countries is implemented according to the regulations set up by the immigration bill; and (ii) the policy as well as the change in policy itself do not display any supply effect. If the latter assumption is violated, it is impossible to disentangle the double-selection process described above. In order to be confident that this assumption is not violated, one has to consider immigration cohorts which entered the countries as directly as possible before and directly after the policy change. The analysis then proceeds in the following steps and is illustrated in Figure 1.
In a first step one compares the earnings position of a specific immigration cohort in country $A$ (i.e. the performance of immigrants entering the country during a given period) relative to comparable natives (i.e. conditional on observable individual characteristics) a sufficiently long time period after the migrant cohort entered the country (five or ten years, say). This yields the relative earnings position of the first cohort in country $A$. The comparison itself could be done in an earnings regression framework or, alternatively, by matching methods. Similarly, one has to calculate the the relative earnings position for the first immigrant cohort in country $B$ by the same econometric procedure.

Then, in a second step, one calculates the first difference, i.e. the difference in the relative earnings positions of the two migrant cohorts before the policy change. This might yield, for instance, a higher earnings disadvantage for migrants in country $B$ then for those in country $A$. This case is graphically depicted in Figure 1. This higher earnings disadvantage for migrants in country $B$ might be the result of a difference in the supply of migrants between both countries or of differences in the institutional settings of the labor markets in determining the assimilation process of the migrant cohorts or it might be the result of an, on average, less successful native comparison group in country $A$. Moreover, and this is the decisive point for the evaluation of immigration policy, this first difference in relative earnings positions might also reflect a more successful selection by the specific regulations laid down in the immigration policy. Therefore, from this (cross-sectional) comparison of the value of the outcome measure it is not possible to isolate a causal effect of a specific policy regime.

Before proceeding with the description of the proposed difference-in-differences approach, it is illustrative at this point to reconsider the received literature again. This reconsideration is restricted to those contributions utilizing individual data on immigrants. Duleep and Regets (1996), for instance, compare the relative earnings performance of two different migrant groups in the US over time in order to analyze if skill-based migrants are more successful than kinship migrants. The authors argue that the catching-up process of kinship migrants is faster than that of skill-based migrants and that the difference in policy plays no substantial role in the long run. The implicit counterfactual question of their analysis, however, is: What would have happened to the relative earnings performance of skill-based migrants (conditional on observables) if they had not been admitted to the US on the basis of their skills? The answer and, therefore, identification assumption of Duleep and Regets (1996) is: The relative performance of this migrant group would have been equal to the relative performance of kinship migrants in the US.

This does, however, not provide a convincing answer to the question whether the skill-based selection mechanism is superior to the kinship immigration regime. From this cross-sectional analysis, i.e. from the difference
in relative earnings positions of both migrant groups after a sufficiently long adaption process, it is impossible to disentangle the double-selection process in immigrant admission.

From the perspective of evaluating whether the Canadian immigration policy has been successful or not, the analysis of BORJAS (1993) compares the upon-entry differences in human capital and earnings of skill-based migrants to the US and Canada. By decomposing these differences according to the country of origin of the immigrants, he concludes that that the Canadian point system is not able to attract more skilled immigrants from a given country of origin. In principle the same argument applies to the analyses of BORJAS (1993) and DULEEP AND REGETS (1996). In both contributions the effect of the first (self-) selection process could only be discriminated from the second, i.e. the selection process due to immigration policy, if one is willing to assume that the supply to both countries is equal in terms of observable as well as unobservable characteristics. Furthermore, for the case of BORJAS' (1993) analysis, it is not really decisive where the immigrants come from but if the policy regime was able to select those from the pool of potential migrants who are the most successful. The analyses of DULEEP AND REGETS (1996) and BORJAS (1993), therefore, demonstrate the importance of a convincing comparison situation for the assessment of the effects of immigration policy.\footnote{The analysis of WRIGHT AND MAXIM (1993) is comparable to that of BORJAS (1993). Furthermore, CHISWICK AND MILLER (1992) study determinants of unemployment conditional on immigrant group and other explanatory factors. This paper is comparable to DULEEP AND REGETS (1996) despite a different outcome measure. Therefore, the same arguments apply for these contributions.}

Consequently, to be able to disentangle both selection effects and, therefore, isolate the causal effect of the immigration policy one essentially needs a regime switch in one of the countries under consideration. In the example at hand, assume that country $B$ changes its policy e.g. towards a more skill-based selection regime. Thus, directly after the regime change in country $B$ there are two new immigrant cohorts entering both countries. For those two cohorts one again has to calculate the relative earnings positions a sufficiently long time period after they have immigrated. This yields the \textit{second difference}, i.e. the difference in the earnings position of migrants in $A$ and $B$ relative to comparable natives \textit{after} the regime change in country $B$ has occurred.

Finally, the treatment parameter, i.e. the mean effect of the regime change on the relative earnings position of immigrants to country $B$ under the new policy regime, is then the \textit{difference} between the first difference and the second difference, i.e. the difference-in-differences.
This difference-in-differences can be interpreted as the causal effect of the change in immigration policy if and only if the following assumptions \textit{jointly} hold:

(i) If there were a difference in the supply of migrants to both countries before the regime change, then the effect of this supply difference on the relative earnings position of immigrants stayed constant over time.

(ii) Similarly, if there were a difference in the institutional settings on the labor markets of both countries before the regime change, then the effect of this institutional difference on the relative earnings position of immigrants stayed constant over time.

(iii) The structure of the earnings performance of the native comparison group in both countries did not change over time.

The central identification assumption of this approach and, thus, the answer for the counterfactual question raised above is then: The difference in the relative labor market performance of the two immigrant cohorts in both countries would have remained constant if the policy regime operating in country $B$ had not changed. If assumptions (i) to (iii) hold, all the differences between both countries which influence the value of the outcome measure and which are \textit{not} related to immigration policy, including potential differences in the supply of migrants, will be netted out by the double differencing approach.

Note that assumption (i) does not necessarily mean that the second immigrant cohort in country $A$ has to have the same relative earnings position after $T$ years in the destination country as the first immigrant cohort. It is possible that there is a change in supply towards more successful migrants (as indicated in \textbf{Figure 1}). However, for the identification assumption to be valid, this “quality” change must be independent of the immigration policy.
regime itself and it must be accompanied by a proportional change in country \( B \) as well. Otherwise, one would fallaciously attribute the effect of this change in supply to the regime switch.

Moreover, assumption (iii) secures that there is no secular trend in the earnings performance of the native comparison group over time in one of the countries rendering the relative earnings position of one migrant cohort better or worse. Clearly, these assumptions are very strong and might be violated easily. However, without these assumptions the difference-in-differences identification strategy is not able to disentangle the effect of the policy regime from the supply effect, i.e. to discriminate between the two selection mechanisms.

Finally, since the relative labor market position of a migrant cohort, typically measured in an earnings-regression framework, decisively depends on the rate of growth in earnings conditional on the years since migrant, i.e. the slope of the earnings function with respect to the duration of residence of the migrant cohort in the country, the difference-in-differences analysis might react sensitively to the chosen evaluation points. It is, therefore, suggested to check the sensitivity of the results by choosing different time periods for the adaption process of immigrants.

4 Conclusions

From the perspective of a country like Germany, serving as a potential destination for people willing to emigrate from their country of origin, a rational regulation of immigration is of central concern for future economic prospects. An ageing society with its consequences for the social security system, an increasing demand for high-skilled labor as well as the prevention of a massive inflow of illegal immigrants will inevitably move the issue of the “best” immigration policy into the center of attention. Unfortunately, economic research on this question has not been able to provide a completely convincing answer. This paper has outlined a conceptual framework for the assessment of the effect of a specific immigration policy by discussing the necessary elements of such a formal evaluation study. It clarified the need to invoke identification assumptions which have to be assumed to hold \( a \text{ priori} \). From this discussion it should have become transparent that it is a conceptually involved task to define an adequate outcome measure and to construct a convincing comparison situation for the unobservable counterfactual situation.

The scarce empirical evidence available at the moment suggests that the regulation of immigration focussing exclusively on the selection of migrants according to a country’s current need for specific skills is not sufficient to guarantee that immigrants are successful on the destination countries labor market. Such a policy runs the risk of neglecting important aspects of the long-run determinants of immigrants’ economic success, i.e. the ability to
cope with a changing economic environment. The experience of the “guest worker” migrants, actively selected by the German immigration policy of the 1960’s may serve as an example in this context. In the first years after their arrival these immigrants were employed and experienced a modest but positive earnings growth (see Schmidt (1997)) compared to similar natives. Their situation, however, has probably changed drastically in the 1980’s when unemployment figures of this immigrant group rose substantially and remained high during the 1990’s (see Fertig and Schmidt (2001)).

The international empirical evidence, furthermore, suggests that a rational and, therefore, foresighted immigration policy should be able to signal reliably that it is in the vital interest of the destination country to admit immigrants with a long-run perspective in the country. It is, therefore, necessary to provide incentives for immigrants to invest into destination-country-specific human capital. In this endeavor it does, for instance, not seem helpful to award work permissions on a temporary basis a priori, as it is done for the so-called “green card” migrants in Germany, or to restrict family reunification tightly as it is discussed for the new German immigration law.

In addition, there seems to be room for an integration policy comprising assistance in acquiring destination country-specific human capital or job search skills to immigrants. It is, therefore, a complementary task of future research in this field, to evaluate the effects of different integration policy measures on the economic success of immigrants.
References


COBB-CLARK, DEBORAH A., CLINTON R. SHIELS and LINDSAY B. LOW-


SCHMIDT, CHRISTOPH M. (2000), Do We Need Social Experiments? mimeo., University of Heidelberg.


VANOURS, JAN C. and JUSTUS VEENMAN (1999), The Netherlands: Old


<table>
<thead>
<tr>
<th>No.</th>
<th>Author(s)</th>
<th>Title</th>
<th>Area</th>
<th>Date</th>
</tr>
</thead>
</table>
| 422  | J. T. Addison  
L. Bellmann  
C. Schnabel  
J. Wagner | The Long Awaited Reform of the German Works Constitution Act         | 6    | 02/02 |
| 423  | E. Feess  
G. Muehlheusser | Transfer Fee Regulations in European Football                       | 1    | 02/02 |
| 424  | F. Büchel  
M. van Ham | Overeducation, Regional Labour Markets and Spatial Flexibility       | 3    | 02/02 |
| 425  | F. Büchel | Successful Apprenticeship-to-Work Transitions: On the Long-Term Change in Significance of the German School-Leaving Certificate | 3    | 02/02 |
| 426  | J. Hartog  
W. P. M. Vijverberg | Do Wages Really Compensate for Risk Aversion and Skewness Affection? | 5    | 02/02 |
| 427  | D. Del Boca | The Effect of Child Care and Part Time Opportunities on Participation and Fertility Decisions in Italy | 6    | 02/02 |
| 428  | D. Del Boca | Mothers, Fathers and Children after Divorce: The Role of Institutions | 6    | 02/02 |
| 429  | S. Anger  
J. Schwarze | Does Future PC Use Determine Our Wages Today? Evidence from German Panel Data | 5    | 02/02 |
| 430  | J. Schwarze  
M. Härpfer | Are People Inequality Averse, and Do They Prefer Redistribution by the State? Evidence From German Longitudinal Data on Life Satisfaction | 3    | 02/02 |
| 431  | M. Fertig  
C. M. Schmidt | The Perception of Foreigners and Jews in Germany - A Structural Analysis of a Large Opinion Survey | 6    | 02/02 |
| 432  | E. Tekin | Employment, Wages, and Alcohol Consumption in Russia: Evidence from Panel Data | 4    | 02/02 |
| 433  | J. D. Angrist  
A. D. Kugler | Protective or Counter-Productive? Labor Market Institutions and the Effect of Immigration on EU Natives | 3    | 02/02 |
| 434  | A. D. Kugler | From Severance Pay to Self-Insurance: Effects of Severance Payments Savings Accounts in Colombia | 4    | 02/02 |
| 435  | G. S. Epstein  
M. E. Ward | Perceived Income, Promotion and Incentive Effects                   | 1    | 02/02 |
| 436  | A. Kunze | The Evolution of the Early Career Gender Wage Gap                   | 1    | 02/02 |
| 437  | M. Fertig | Evaluating Immigration Policy Potentials and Limitations             | 6    | 02/02 |

An updated list of IZA Discussion Papers is available on the center's homepage [www.iza.org](http://www.iza.org).