SUMMARY

Discussions of immigration policy are typically framed in the context of their economic effects in receiving countries, notably labour market and fiscal effects. In this paper, we characterize immigration policy in a richer model where migrants are also a source of cultural externalities stemming from either preferences or the functioning of formal and informal institutions in receiving countries. While in terms of pure economic effects, immigrants do not generally have any more incentives than low-skilled natives to allow for more immigration in the future, this is not the case when accounting for cultural externalities. Therefore, insofar as past immigrants have a voice in affecting future policies, a time-consistent immigration policy entails back-loading; as natives attempt at limiting voice of immigrants in the future, the economic effects of immigration flows as well as the cultural externality they introduce. Furthermore, natives exploit any pre-commitment device to limit immigration flows, e.g. building “walls”, limiting immigrants’ political rights, or accumulating fiscal surpluses.

JEL codes: E61, F22, J61, Z1

—Alberto Bisin and Giulio Zanella
1. INTRODUCTION

Few countries have had consistent immigration policies over long periods of time. Japan, however, is one that has: it has consistently kept immigrants out throughout its history. —Sowell (1996, p. 44)

Surveys of European voters’ attitudes towards immigration reveal large potential support for restrictive immigration policies, with sizable heterogeneity across countries.
Figure 1 illustrates these attitudes using data from the European Social Survey (ESS). Native individuals are asked to what extent they believe that their government should allow ‘few’ or ‘none’ immigrants to come and live there, as opposed to ‘some’ or ‘many’. Sampling weights are applied. For Greece, Italy and Slovakia, this information is not available in the 2014 wave of the ESS, and it is replaced by the most recently available data (2010 for Greece, 2012 for Italy and Slovakia).

While economic analyses suggest that the welfare gains from an ‘open borders’ immigration policy are large (e.g. Freeman, 2006; Clemens, 2011; Kennan, 2013, 2017), popular opposition to these policies may reflect the perceived or actual impact of migrants on the labour market and welfare system of receiving countries, and the difficulty in implementing mechanisms that redistribute the large welfare gains generated by an open immigration policy from the winners (first and foremost not only the immigrants

---

1 For Greece, Italy and Slovakia, this information is not available in the 2014 wave of the ESS, and it is replaced by the most recently available data: 2010 for Greece and 2012 for Italy and Slovakia. For brevity, though, we always refer to year 2014.
themselves, but also employers) to the losers (low-skilled workers, but also users of congestable public welfare services, possibly).\(^2\)

Without disputing this argument, we suggest that the opposition to open borders also reflects negative cultural externalities that immigration imposes on natives in the process of social interaction. Based on this idea, we present a theory of immigration policy where low-skilled and high-skilled natives are affected by immigrants via three channels: labour market outcomes, congestable public goods and cultural externalities.\(^3\) The theory is intentionally simple, and is meant to illustrate in a transparent way the consequences of cultural concerns for our understanding of natives’ aversion to open borders in European countries and elsewhere. We posit that this class of external effects arises from different channels: direct and indirect preferences, and the functioning of informal and formal institutions. The model implies that such cultural externalities induce natives to oppose a possibly welfare-improving open immigration policy even in the absence of any labour market or fiscal impact. Furthermore, the skill heterogeneity of natives and the presence of cultural externalities make immigration policy time-inconsistent, despite individual preferences being not necessarily so. This time-inconsistency induces natives to oppose immigration flows more strongly than they would if the government had commitment over its future immigration policy. Moreover, the time-inconsistency induces a demand for commitment devices by part of natives, i.e. strategies to increase the cost of admitting migrants in the future (or, equivalently, reducing the cost of keeping them out). One such strategy is building immigration ‘walls’ (Brexit is a case in point). Other strategies include limiting immigrants’ political rights and accumulating fiscal surpluses (i.e. a tight fiscal policy).

Our message is that a deeper understanding of the nature and consequences of the cultural externalities generated by immigration helps gaining deeper insights into immigration policy. We convey this message in various steps. Section 2 illustrates the notions of economic and cultural externalities from immigration. In Section 3, we build a model formalizing such economic and cultural concerns, and we use it to characterize immigration policy when the government can commit to a planned immigration sequence (commitment immigration policy), and when it cannot commit (no-commitment, or equilibrium immigration policy). In Section 4, data from the ESS is analysed to gauge at the driving mechanisms embedded in the model. Section 5 concludes.

---

2 Auctions to allocate employment permits to employers and work permits to those immigrants who can contribute the most to economic activity in a specific country is one such mechanism (Peri, 2012).

3 Other channels are possibly relevant, such as the impact of immigration on crime rates (see Bell and Machin, 2013 for a review) or on innovation in the host country (Hunt and Gauthier-Loiselle, 2010). These other channels are not considered here.
2. THE ECONOMIC AND CULTURAL IMPACT OF IMMIGRANTS

2.1. Economic impact

The labour market and fiscal effects of migrants in host countries are elusive. Despite a massive empirical literature, there is no consensus among economists about the impact of immigration on the labour market outcomes of different workers. One view is that immigration to the United States over the past 30 years has reduced the wages of native workers, with larger effects for low-skilled individuals, both in the short and in the long run (e.g. Borjas et al. (1997) and Borjas (2003, 2014) offer a comprehensive summary of this view). A contending view is that the long-run impact of immigrants on the wages of natives is positive both in the United States and the United Kingdom (Ottaviano and Peri (2012), Manacorda et al. (2012) and Card (2009) summarize this alternative view). An intermediate view is that the impact of immigrants on wages is heterogeneous at different points of the wage distribution: negative at lower quantiles, positive at upper quantiles (Dustmann et al., 2013). The disagreement reflects different choices in the setup of the underlying econometric or theoretical models. How one measures the immigration flow, how fast the capital stock adjusts, how one classifies education groups, the labour supply elasticity and the elasticity of substitution between immigrants and natives within the same education group and, of course, the details of the empirical specification, all make an important difference (Card, 2012; Dustmann et al., 2016; Card and Peri, 2016). There is no convergence of opinions on these issues.4

The fiscal effects of immigration are similarly elusive. If the new immigrants contribute to the fiscal system a per-capita amount equivalent to the per-capita public good enjoyed by residents before their arrival, then immigration is fiscally neutral and there is no external effect.5 However, measuring the net fiscal contribution of immigrants raises a number of conceptual and empirical issues, such as how to implement a multi-period, multi-generation measurement, whether tax incidence or tax revenue is the appropriate outcome to look at, the degree to which government services are subject to congestion and the marginal cost of providing government services (MaCurdy et al., 1998; Preston, 2014). A recent overview of existing estimates in the light of some of these issues can be

4 This conclusion is based on studies of labour markets in the United States and the United Kingdom, which are similar in many respects but different from labour markets in continental Europe. Also, the skill composition of the US and the UK labour forces, as well as the composition of the immigration flows resulting from self-selection of immigrants differ from continental Europe. These differences suggest caution when extrapolating results for the United States and the United Kingdom to other labour markets. However, the theoretical and methodological issues arising from studies of the United States or the United Kingdom are general. Moreover, Docquier et al. (2014) provide a consistent set of model-based estimates of the impact of immigration on natives’ wages and employment for a large group of OECD countries: qualitatively, results for the United States and the United Kingdom extend to the bulk of the OECD group, although there is interesting heterogeneity.

5 This neutrality result is based on the assumption that the tax contributions of immigrants are not diverted from public good provision by the government. Whether they are or not is a political economy issue, not an effect of immigration per se.
found in Vargas-Silva (2014) and Preston (2014). Liebig and Mo (2013) provide, to the best of our knowledge, the only existing comparative analysis of the net fiscal impact of immigration for OECD countries, although based on a static accounting framework.

The pattern emerging from this research is that the net fiscal impact of immigrants is positive in some countries and negative in others, but it is small relative to GDP, and essentially zero, on average, across European countries. According to Liebig and Mo (2013), the heterogeneity in the net fiscal impact of immigrants in Europe reflects more the taxes they pay than the benefits they receive in different countries. This observation is consistent with the fact that empirical analysis does not provide strong support for the ‘welfare magnet’ hypothesis (Borjas, 1999).

2.2. Cultural impact

Natives may care about the cultural composition of their country for a variety of reasons, and an inflow of immigrants carrying a distinctive cultural identity different from the natives’ increases the cultural heterogeneity of the country and generates a cultural externality. Similarly, immigrants care about this same cultural composition and they are also affected by the arrival of other immigrants. Several such reasons are emphasized in the literature, as reviewed in Bisin and Verdier (2011).

First, the direct and indirect preference effects. A direct preference effect on natives occurs because of pure cultural intolerance. For example, the utility of a native may be decreasing in cultural heterogeneity because of a pure distaste for contact with different identities associated with different behavioural prescriptions (Akerlof and Kranton, 2000), such as when a native feels that immigrants threaten the cultural identity of the host country. An indirect preference effect, instead, occurs when native parents are biased towards the native culture in the process of building their offspring’s identity due to imperfect empathy, whereby native parents evaluate their offspring’s outcomes through their own preferences (Bisin and Verdier, 1998). This form of paternalistic altruism induces preferences over the cultural composition of the community, because such composition has a bearing on the process of cultural socialization of children and, possibly,

---

6 The underlying heterogeneity indicates that immigration is more costly for some European countries than others, suggesting that transfers within Europe neutralizing the net fiscal impact of immigrants are possible, in principle. The tradable immigration quota scheme analysed by Fernández-Huertas Moraga and Rapoport (2014) is an example of how such transfers can be implemented in practice.

7 Reviewing the evidence about welfare benefits as driving self-selection of immigrants, Preston (2014) concludes that ‘while several studies have found evidence linking welfare to immigrant selection in particular cases, this is only one among several factors influencing migration decisions and no uniform picture emerges of relative welfare dependency among immigrants.’ (p. F576). That is, there seems to be no important or at least no consistent self-selection of immigrants in response to the different generosity of welfare systems across different locations. Moreover, like for the labour market impact of immigrants, mechanisms exist even for a single country to internalize the externalities arising from congestion of public goods. For instance, the eligibility of immigrants for several means-tested public programmes is restricted in the United States (Wasem, 2014).
grandchildren in a context in which the socialization technology available to parents takes as inputs their own effort and societal effects. For example, immigration affects the cultural composition of schools and neighbourhoods, in a context where peers matter in the development of a child’s identity. This ‘oblique’ socialization mechanism implies a negative externality for parents of a certain cultural type when their children are exposed to a different type. Another example is the marriage market, whose composition is affected by the presence of immigrants. Heterogamous marriages (as opposed to homogamous ones) are associated with different socialization technologies because in mixed marriages the socialization efforts of parents work in opposite directions (Bisin and Verdier, 2000). This ‘vertical’ socialization mechanism implies a negative externality if it increases the likelihood that one’s children end up in a heterogamous marriage. Previous immigrants may instead value positively the arrival of more immigrants of the same type (e.g. ethnic or religious group) because the presence of a larger community sharing their own cultural identity facilitates the process of socialization of their own children via both oblique and vertical socialization. The cultural externality is positive for them.8

Second, the informal and formal institutional effects. An efficient institutional system optimally trades off informal and formal institutions. Informal institutions (e.g. reputational concerns facilitating the enforcement of contracts or virtuous forms of social control preventing crime) are relatively inexpensive but require civic capital, i.e. ‘persistent and shared beliefs and values that help a group overcome the free rider problem in the pursuit of socially valuable activities’ (Guiso et al., 2011, p. 419). Formal institutions (e.g. the political and judicial systems) are relatively expensive but they, too, require civic capital to work effectively (Putnam, 1993). It has been argued that an increase in ethnic and cultural heterogeneity, like the increase that would follow from a large influx of immigrants of a different ethnic group from the majority, may lead to a reduction in the stock of civic capital because trust, ability to cooperate and shared norms are scarcer in heterogeneous communities (Alesina and La Ferrara, 2002, 2005). Like for preference effects, this cultural externality arising from institutional effects is negative for the natives, who own the stock of civic capital at the baseline, pre-immigration state but is positive for immigrants, who build more of their own civic capital as more immigrants of the same type enter the economy.

3. MODEL

Taking stock of the labour market, fiscal and cultural aspects of immigration discussed above, we now introduce a simple theoretical model whose objective is to frame a more formal discussion of the effects of immigration and to provide a tool for the study of

---

8 Whether a larger or smaller cultural group of one’s own type increases or decreases parents’ socialization effort depends on whether the socialization technology exhibits cultural substitution (cultural distinction) or cultural complementarity, i.e. whether minorities, other things equal, socialize more or less intensely their children. As a consequence, these mechanisms depend, among the other things, on the level of segregation and discrimination of immigrants.
immigration policies when cultural externalities matter. While the model can be easily extended to perform quantitative exercises, we intentionally present it here in its simplest, bare-bone form to strip down the analysis to its fundamental theoretical components.

3.1. Set-up

The economy is populated by a group of natives who live for two periods, \( t = 1, 2 \), and who are either high-skilled native workers (\( h \)) or low-skilled native workers (\( l \)), as well as by a group of immigrant workers. Immigrants are (and remain, by assumption) low-skilled workers in the labour market of the host country. The native population is constant over time and is normalized to 1, so that \( L_h \) native workers are high-skilled and \( L_l = 1 - L_h \) are not. At time \( t \), a flow of \( m_t \) immigrants (expressed as a fraction of the constant native population) are admitted into the economy. Let \( q_t = \sum_{t=1}^{t} m_t \) denote the stock of immigrants at time \( t \), i.e. the immigrant/native ratio in the country.

Border enforcement is costly, i.e. it is costly to keep out migrants wishing to enter the country. For instance, a restrictive immigration policy may result in attempts to enter the economy in an unauthorized way, which is costly to contain. A stock of \( M_t \) migrants are ready to enter the economy, a number that we take as given. Let \( \zeta(\varphi_t) \) denote the associated cost function, i.e. the minimum cost of producing an inflow of \( m_t \) migrants for \( t = 1, 2 \).\(^9\) We assume \( \zeta'(\varphi_t) < 0 \) and \( \zeta''(\varphi_t) > 0 \). That is, the more immigrants the country admits, the less costly border enforcement is. In the extreme case of open borders, it is \( \zeta(M_t) = 0 \). The marginal cost of border protection is assumed to be increasing, and hence the convexity of \( \zeta(\varphi_t) \).

The labour market is characterized by inelastic labour supply and elastic labour demand functions for high-skilled and low-skilled labour. Denote by \( w_h^t \) and \( w_l^t \) the high-skilled and low-skilled wages, respectively. At equilibrium, for given \( L_h^t \), \( w_h^t = w_h(q_t) \) and \( w_l^t = w_l(q_t) \). We also assume that low-skilled and high-skilled workers are complements in production, so that \( \frac{\partial w_h(q_t)}{\partial q_t} > 0 \) and \( \frac{\partial w_l(q_t)}{\partial q_t} < 0 \), \( t = 1, 2 \). These assumptions embed in our model the evidence that immigration affects natives’ wages positively at the top of the wage distribution and negatively at the bottom (Dustmann et al., 2013). Therefore, as far as labour market effects are concerned, high-skilled natives would like more immigration, while low-skilled natives would like less.

Natives and immigrants have identical preferences over private consumption, \( c_t \). However, low-skilled workers (natives and immigrants alike) value a government-provided public good, \( g_t \), more than high-skilled workers. The latter, possibly, do not value the public good at all because they are already consuming privately provided substitutes. The public good is provided by means of exogenous public expenditure \( \gamma_t \).

---

\(^9\) We assume that this cost depends on the total migration stock, \( q_2 = m_1 + m_2 \), independently of the distribution of the flows over time. This is for simplicity and easy to relax.
financed by lump-sum taxes \( \tau \), on both natives and immigrants. The government budget is balanced in each period \( t \), i.e. \( \gamma_t = \tau_t \). Preferences over goods are represented by a strongly monotonically increasing and strictly concave function \( u^\theta(c_t^\theta, g_t) \), with \( \theta = h, l \) and \( \frac{\partial^2 u}{\partial g^2} > \frac{\partial^2 u}{\partial c^2} > 0 \), which agents maximize in each period \( t \) subject to the budget constraint and a public good provision constraint:

\[
c_t^\theta = w_t^\theta - \tau_t, \quad \theta = h, l;
g_t = \gamma_t - g(q_t).
\]

Function \( g(q_t) \) is increasing, with \( g(0) = 0 \), and captures a congestion effect arising, for instance, from decreasing returns in the production of the public good. Therefore, negative fiscal effects of immigration are embedded in our model, and as far as this fiscal externality is concerned, high-skilled natives are less averse to immigration (and possibly indifferent) than their low-skilled counterpart.

As for cultural identity, natives are characterized by a cultural trait that is different from the cultural trait of immigrants at arrival. The latter is denoted by \( i \). Because immigrants are low skilled, it is \( u'(c_t^i, g_t) = u'(c_t^l, g_t) \), i.e. immigrants carrying cultural trait \( i \) value private and public consumption like low-skilled natives. These traits do not change over time, but immigrants can (choose to) assimilate to the culture of the host country, in which case they acquire the natives’ cultural trait, i.e. they become indistinguishable from low-skilled natives. Let \( p_t \) denote the stock of non-assimilated immigrants. As immigrants are not assimilated at arrival, \( p_1 = q_1 = m_1 \). The cultural dynamics from period \( t=1 \) to \( t=2 \) is represented by the population dynamics of the distribution of cultural traits:

\[
p_2 - p_1 = f(p_1) + m_2,
\]

where the map \( f(p_1) \) represents cultural identity formation. Because we have assumed that immigrants may assimilate to the host culture but natives keep their cultural trait, \( f(p_1) \leq 0 \) represents the (negative of the) assimilation rate of immigrants. That is, of the \( m_1 \) immigrants admitted at \( t=1 \), \(-f(m_1)\) are culturally assimilated to the host country by \( t=2 \) while \( m_1 + f(m_1) \) keep their original cultural identity. Several microfoundations of Equation 1 (see Bisin and Verdier, 2011) suggest the following function form for \( f(p) \):

\[
f(p_1) = p_1(1 - p_1)d(p_1),
\]

where \( d(p_1) \leq 0 \) captures the socialization effort of immigrant parents, and is assumed to satisfy \( d'(p_1) < 0 \), i.e. ‘cultural substitution’. Roughly speaking, this means that immigrant parents’ effort in transmitting cultural identity \( i \) to their children (‘direct
socialization’, via family interactions) decreases with the frequency of cultural trait $i$ in the community where the children grow up (‘horizontal socialization’, via social interactions). In the special case in which assimilation is effectively impossible (or not desired), it is $f(p) = 0$ and so $p_t = q_t$.

Cultural externalities are captured by a component of each agent’s preferences which depends on his/her cultural profile and skill group. Presumably, high-skilled natives have limited social contact with immigrants relative to the low-skilled, and so they are less subject to the cultural externality. The cultural component of preferences is thus represented by an additive indirect utility term $r^\theta(p_t)$, with $\theta = h, l, i$, which embeds in a reduced-form way all of the cultural externalities discussed in Section 2.2. The discussion there and the assumption that high-skilled natives have less social contacts with immigrants than low-skilled ones imply that $r^\theta(p_t)$ decreases in $p_t$ for $\theta = h, l$, while it increases in $p_t$ for $\theta = i$, with $|\frac{\partial \psi}{\partial q_t}| > |\frac{\partial \varphi}{\partial p_t}|$. In other words, the cultural externality of immigration is negative for natives but is positive for non-assimilated past immigrants. Denoting by $v^\theta(q_t) + r^\theta(p_t)$, the value function resulting from type $\theta$ agents’ maximization subject to the budget and public good provision constraints, preferences over the cultural composition of the country are represented by:

$$v^\theta(q_t) + r^\theta(p_t), \quad \theta = h, l, i. \tag{3}$$

Notice that the component of preferences reflecting labour market outcomes and the public good, $v^\theta(q_t)$, is the same for low-skilled natives and immigrants, contrary to their different evaluation of the cultural composition of the economy, $r^\theta(p_t)$, unless the immigrants in question are culturally assimilated. High-skilled natives and non-assimilated immigrants, instead, differ in the evaluation of all of three impacts of additional immigration: wages, public good and culture.

3.2. Social welfare

Without explicitly specifying an institutional (e.g. voting) mechanism which delineates and implements policies, we assume that immigration policies are the result of the maximization of a government’s objective function which takes into account the welfare of both natives and immigrants. Immigrants’ welfare, in particular, is taken into account by the government for political, economic, social or humanitarian reasons that are not specified here. Then social welfare weights $\psi^\theta_t$, $\theta = h, l, i$, $t = 1, 2$ are monotonically

---

11 Bisin and Verdier (2011) summarize the evidence regarding cultural substitution.

12 Iceland and Scopilliti (2008) have documented that the residential segregation of immigrants in the United States is higher than for minorities. Even within neighbourhoods, Petermann and Schönwälter (2014) show some evidence (Table 1 therein) that the frequency of social interactions between immigrants and previous residents decreases with the income and education levels of the latter.

13 The separability in Equation (3) simplifies the exposition, but it is not driving the results derived below.
increasing functions of population shares at the beginning of the period (prior to immigration flows):

\[
\psi_1^h = \psi^h(L^h), \quad \psi_1^l = \psi^l(1 - L^h), \quad \psi_1^i = 0;
\]

\[
\psi_2^h = \psi^h\left(\frac{L^h}{1 + m_1}\right), \quad \psi_2^l = \psi^l\left(\frac{1 - L^h}{1 + m_1}\right), \quad \psi_2^i = \psi^i\left(\frac{m_1}{1 + m_1}\right).
\]

These weights are normalized so that \(\sum_t \psi_t^i = 1\), for \(t = 1, 2\). Note that \(\psi_2^l\) and \(\psi_2^h\) decrease with \(m_1\), while \(\psi_2^i\) increases in this quantity. Therefore, since these sum up to 1, it must be that \(\psi_2^l + \psi_2^i\) increases with \(m_1\): immigration at \(t = 1\) reduces the ‘power’ (as represented by the weights in the government’s objective) of high-skilled natives and increases the ‘power’ of low-skilled workers, including the immigrants themselves, at \(t = 2\).

We distinguish two social welfare choice problems, with and without commitment, and we emphasize that our analysis is positive: the social welfare choice problem with no-commitment is the policy choice equilibrium outcome in our economy (again, the policy-maker cares about groups in proportion to their incidence for political, economic or social reasons captured by the social welfare weights). The commitment problem is instead a useful benchmark, to identify the effects of time-inconsistency at equilibrium.

The commitment immigration policy is a sequence \(\{m_1^c, m_2^c\}\) which maximizes the social welfare function under commitment, i.e. when the immigration flows both at \(t = 1\) and \(t = 2\) are chosen at \(t = 1\) and maintained. This is the solution to

\[
\max_{m_1, m_2} \sum_{\theta = h, l, i} \sum_{t = 1, 2} \psi_t^\theta[y_t^\theta(q_t) + r_t^\theta(p_t)] - \varphi(q_2)
\]

s.t. \(q_1 = p_1 = m_1\); \(q_2 = m_1 + m_2\); \(p_2 = p_1 + f(m_1) + m_2\).

The equilibrium (or no-commitment) immigration policy, instead, is a sequence \(\{m_1^w, m_2^w\}\) which maximizes the social welfare function without commitment. In this case the immigration flow at \(t = 2\) is chosen at \(t = 2\), after the flow at \(t = 1\) has been chosen. At \(t = 1\), the flow at \(t = 2\) is perfectly anticipated as a function of the choice at \(t = 1\). This policy solves

\[
\max_{m_1, m_2} \sum_{\theta = h, l, i} \sum_{t = 1, 2} \psi_t^\theta[y_t^\theta(q_t) + r_t^\theta(p_t)] - \varphi(q_2)
\]

s.t. \(m_2 \in \arg\max \sum_{\theta = h, l, i} \psi_2^\theta[y_2^\theta(q_2) + r_2^\theta(p_2)] - \varphi(q_2)\);

\[
q_1 = p_1 = m_1; \quad q_2 = m_1 + m_2; \quad p_2 = p_1 + f(m_1) + m_2.
\]

The difference between the two problems originates from constraint (6), which captures the fact that the immigration policy at \(t = 2\) is chosen in fact at \(t = 2\). Let
\[ V_1(m_1) = \sum_{\theta=h,l,i} \psi_1^\theta[v_1^\theta(m_1) + r^\theta(m_1)], \]
\[ V_2(q_2, p_2) = \sum_{\theta=h,l,i} \psi_1^\theta[v_2^\theta(q_2) + r^\theta(p_2)], \]
\[ H_2(q_2, p_2) = \sum_{\theta=h,l,i} \psi_2^\theta[v_2^\theta(q_2) + r^\theta(p_2)]. \]

We impose stringent but standard assumptions on the concavity–convexity of the maps \( w^\theta(q), g^\theta(q), \) and \( r^\theta(q) \) so that the commitment problem can be guaranteed to be convex and first-order conditions are necessary and sufficient for a unique maximum. The no-commitment problem, instead, is generally not convex. But can we impose Inada conditions guaranteeing that an interior maximum satisfying the first-order conditions exists.14 The first-order conditions for the commitment problem, after exploiting the Envelope condition, are reduced to the constraints in Equation (7) and

\[ \frac{dV_1(m_1)}{dm_1} + \frac{\partial V_2(q_2, p_2)}{\partial m_1} = \frac{\partial z(q_2)}{\partial m_1}, \]  
(8)
\[ \frac{\partial V_2(q_2, p_2)}{\partial m_2} = \frac{\partial z(q_2)}{\partial m_2}. \]  
(9)

The first-order conditions for the no-commitment problem, instead, are15

\[ \frac{dV_1(m_1)}{dm_1} + \frac{\partial V_2(q_2, p_2)}{\partial m_1} + \frac{\partial V_2(q_2, p_2) dm_2}{dm_1} = \frac{\partial z(q_2)}{\partial m_1} + \frac{\partial z(q_2)}{\partial m_2} \frac{dm_2}{dm_1}, \]  
(10)
\[ \frac{\partial H_2(q_2, p_2)}{\partial m_2} = \frac{\partial z(q_2)}{\partial m_2}. \]  
(11)

where \( \frac{dm_2}{dm_1} \) is obtained by the Implicit Function Theorem on (11), and the constraints in (7). We are now ready to characterize the immigration policy in this model.

A fundamental property of the immigration policy choice is that it is time-inconsistent. At equilibrium (without commitment), social welfare at \( t = 2 \) weights the preferences of immigrants, while only the natives’ preferences enter social welfare at \( t = 1 \). Therefore, time-consistent immigration policy entails the natives, when choosing \( m_1 \) at \( t = 1 \), anticipating the subsequent policy choice \( m_2 \). In other words, when determining the immigration flow at \( t = 1 \) the natives, we take into account that in the future immigrants will contribute to the determination of immigration policy, via their positive weight in the social welfare function (i.e. their ‘political power’). This distorts the immigration policy, at equilibrium, with respect to the policy which would have been chosen under commitment. More formally:

14 For instance, \( \lim_{q \to 0} \frac{\partial z(q)}{dq} = \infty \) and \( \lim_{q \to 0} \frac{dV_2(q, p)}{dq} < 0 \), for \( \theta = h, l, m \), would suffice.

15 Note that the Envelope condition does not hold in this case.
Proposition 1. The social welfare choice problem is generically\textsuperscript{16} time-inconsistent; that is, the no-commitment solution for immigration flows, \( \{m_{1}^{nc}, m_{2}^{nc}\} \), is generically distinct from the commitment solution, \( \{m_{1}^{c}, m_{2}^{c}\} \).

**Proof.** See Appendix.

We now turn to characterize the no-commitment solution, \( \{m_{1}^{nc}, m_{2}^{nc}\} \), and to compare it with the commitment solution, \( \{m_{1}^{c}, m_{2}^{c}\} \). We consider two different special environments: first, one in which the cultural externality is assumed away, and another in which assimilation of immigrants to the natives’ culture is not possible.

Consider first the environment in which cultural identity is not a source of externality; that is, \( r(p) = 0, \theta = h, l, i \).

Proposition 2. Absent cultural externalities, \( r(p) = 0, \theta = h, l, i \), the total immigration flow is smaller at equilibrium (with no-commitment) than with commitment:

\[
m_{1}^{nc} + m_{2}^{nc} < m_{1}^{c} + m_{2}^{c}.
\]

Moreover, the immigration flow is smaller at equilibrium (with no-commitment) than with commitment in the first period:

\[
m_{1}^{nc} < m_{1}^{c}.
\]

**Proof.** See Appendix.

To understand this result, notice that in the absence of a cultural component of preferences it is still the case that social welfare at \( t = 2 \) weights the preferences of immigrants, while only natives’ preferences enter social welfare at \( t = 1 \). Therefore, time-inconsistency is not a consequence of cultural externalities per se. However, the welfare evaluation of the effects of future immigration would be aligned between low-skilled natives and immigrants, because in the absence of cultural externalities, cultural traits are irrelevant and immigrants are effectively replicas of native low-skilled workers: formally, \( v(q) = v(q) \) and \( g(q) = g(q) \). Thus, in this environment, the time-inconsistency of social welfare is due to labour market and fiscal effects only: low-skilled workers (natives and immigrants alike) want to limit future immigration to avoid negative economic externalities. This is a consequence of the two-fold fact that (a) new immigrants at \( t = 2 \) affect the labour market through the wage map \( w(q_{1}) \), which is the same for natives and immigrants who have previously (at time \( t = 1 \)) entered the economy; (b) congestion of the public good also affects natives and immigrants who entered at time \( t = 1 \) in the same way under our budget balance assumption, \( \gamma_{t} = \tau_{p} \) for \( t = 1, 2 \).

\textsuperscript{16} Genericity is to be intended in the space of suitably parametrized utility functions.
At \( t = 2 \), therefore, social welfare with no-commitment will weight the low-skilled more than social welfare with commitment, thereby choosing a lower immigrant flow at \( t = 2 \) than with commitment. Anticipating this effect, at \( t = 1 \), the social welfare problem with no-commitment will reduce the immigration flow \( m_1 \) with respect to the commitment level, to reduce the weight of social welfare on low-skilled and immigrants at \( t = 2 \), effectively accommodating in part the preferences of high-skill workers who will pay the cost of the time-inconsistency at \( t = 2 \). This limits the reduction in the no-commitment immigration flow at \( t = 2 \), without overturning it.

Consider now an environment where culture is a source of externality \( r^\theta(p), \theta = h, l, i \). Assume first that assimilation of immigrants to the natives' culture never takes place, perhaps because it is never desired. That is, immigrants keep their original cultural trait \( i \). In this case it is \( p_i = q_i \). Bisin et al. (2011) and Algan et al. (2012) provide evidence that indeed cultural integration of immigrants is very slow, so that this assumption provides a useful benchmark.

It can easily be shown that, allowing for culture with no assimilation, the commitment solution \( m^c_1 \) does not change, while \( m^c_2 \) is reduced to account for the negative effects of the cultural externality on both high- and low-skilled natives. As for the comparison between the commitment and no-commitment solutions:

**Proposition 3.** Allowing for cultural externality \( r^\theta(p), \theta = h, l, i \), but with no-assimilation, \( p_i = q_i \), the total immigration flow is larger at equilibrium (with no-commitment) than with commitment if the cultural externality is sufficiently large for past immigrants:

\[
m_{1c} + m_{2c} > m_1^c + m_2^c \quad \text{if} \quad \frac{dr^\theta(q)}{dm_2} \quad \text{is sufficiently large}
\]

However, the immigration flow is smaller at equilibrium (with no-commitment) than with commitment in the first period:

\[
m_{1c} < m_1^c.
\]

**Proof.** See Appendix.

Allowing for the cultural externality introduces a component of the preferences of immigrants which values future immigration flows positively. The time-inconsistent social welfare function generally weights the immigrants preferences and hence internalizes these preferences. On the other hand, in terms of their effects on wages and on the public good, immigrants of course, value immigration flows negatively. The relative strength of these components of the preferences of immigrants determines whether at equilibrium (with no-commitment) the immigration flow is larger or smaller than with commitment. In either case, however, the social welfare choice anticipates this effect and in part limits it by reducing the immigrant flow at \( t = 1 \), thereby lowering the weight of immigrants at \( t = 2 \).
Consider finally the general case, where a fraction of the immigrants at time $t = 1$ assimilate to the natives at $t = 2$. Suppose $m_1 < \frac{1}{2}$, so that immigrants are effectively a minority, at least at $t = 1$. In this case the effect of the immigrant flow at $t = 1$ on wages and on the public good is unchanged, but it is more limited on the cultural externality, that is, $\frac{dm_1}{dm_2} + f(m_1) < 1$. Indeed, in this environment, $m_1^\ast$ is larger than with no assimilation, for any $m_1^\ast$ and so is $m_2^\ast$. As long as assimilation is relatively contained, we can show that the general comparison between the flows at equilibrium (with no-commitment) and with commitment is not changed. In particular, it is still the case that $m_{nc}^\ast$ is greater or smaller than $m_2^\ast$, depending on $\frac{\partial f(q)}{\partial m}$, and that $m_{nc}^\ast < m_1^\ast$.

### 3.3. Commitment strategies

In the presence of negative cultural externalities, the time-inconsistency of the social welfare function induces at a hypothetical time $t = 0$ a demand, on the part of natives, for commitment strategies, i.e. choices and distortions which would increase perceived cost of immigration at time $t = 1$ and $t = 2$. While in the bare model we are using no such strategy is available, we can envision several interesting extensions which would allow for them. We offer the reader three examples.

First, investing in irreversible border protection devices. For instance, the natives may ask the government to tax them and build a wall around the border at time $t = 0$. This is tantamount to a downward shift in the marginal cost of border protection, $\alpha'(q_2)$. At that point, it is relatively inexpensive to keep immigrants out, and so the government would choose, at the no-commitment equilibrium, small immigration flows relative to what it would have chosen had the wall not been built. Therefore, in the presence of this commitment device, border policy at equilibrium would be more stringent than under commitment.

Second, reducing the welfare weight on immigrants. For instance, the natives may ask the government to limit immigrants’ political rights (such as voting rights) at time $t = 0$. This would result in a downward shift in welfare weight function $\psi_2$. At that point, the government admits fewer new immigrants at $t = 2$ (i.e. it chooses a low $m_2$) than it would have done in the case of a larger welfare weight on past immigrants. Therefore, in the presence of this commitment device, immigrants’ political rights at equilibrium would be more limited than under commitment.

Third, engaging in a tight fiscal policy. To see why this is a possible commitment device, let’s relax the budget balance assumption in every period, so that the government now can run a deficit or a surplus. Assume that $\gamma_t = \gamma$ exogenously, $t = 1, 2,$ and

---

17 Formally, consider $p + f(p)$ to be a local perturbation of the identity map, $I(p) = p$ which satisfies the assumptions of cultural substitution: $d(p) \leq 0$ and $d'(p) < 0$. This is the case, for instance, if $-\epsilon \leq d(p)$, for a small $\epsilon$.

18 A ‘wall’ is a metaphor for any costly and hard to reverse choice making borders less permeable. In this sense, the 2016 ‘Brexit’ was like building a wall.
that natives and immigrants are taxed homogeneously. Consider fiscal policies in which \( \tau_1 \), the lump-sum tax to be paid by natives and immigrants alike at time \( t = 1 \), is chosen at time \( t = 0 \); while \( \tau_2 \) must satisfy intertemporal budget balance, that is, 
\[
(\gamma - \tau_1)(1 + m_1) + (\gamma - \tau_2)(1 + m_1 + m_2) = 0.
\]
When \( \gamma < \tau_1 \), that is, fiscal policy is characterized by a fiscal surplus at time \( t = 1 \), is it more costly, for both natives and immigrants entered at \( t = 1 \), to admit new immigrants at \( t = 2 \). This is because the new immigrants will participate in sharing the proceeds of the surplus at \( t = 1 \) by facing lower taxes \( \tau_2 \), at \( t = 2 \). In other words, a fiscal surplus is a commitment device which the natives could use to restrict the incentives to admit immigrants in the future. This commitment device is costly, in that the fiscal surplus created at time \( t = 1 \) is shared with the new immigrants at \( t = 2 \). On the other hand, when \( \gamma > \tau_1 \), that is, fiscal policy is characterized by a fiscal deficit at \( t = 1 \), it is less costly, for both natives and immigrants entered at \( t = 1 \), to admit new immigrants at \( t = 2 \). This is because the new immigrants will have to contribute to the repayment of the debt contracted at \( t = 1 \) by means of higher taxes \( \tau_2 \), at \( t = 2 \). In other words, a fiscal deficit reduces the fiscal burden for natives and immigrants entered at \( t = 1 \); but, in the presence of time inconsistency, it represents a cost for the natives because it induces stronger incentives to admit immigrants in the future. Therefore, in the presence of this commitment device, fiscal policy at equilibrium would entail smaller budget deficits than under commitment.

In sum, the model shows that the presence of economic and cultural externalities makes immigration policy time-inconsistent, so that the chosen immigration flow in the early period is smaller than it would be under commitment. Moreover, natives would exploit commitment devices, when available, that further limit the immigration flow or immigrants’ voice in future immigration policy. Contrary to economic externalities, which favour high-skilled natives and so make them favourable to more immigration, cultural externalities reconcile high- and low-skill natives: the cultural motive makes both these groups averse to more immigration – although more strongly so the low-skilled. Although at this level of analysis we cannot test the theory, we provide in the next section evidence that is consistent with it.

4. EVIDENCE

The ESS allows us to gauge at the key driving mechanisms we have embedded in the model, namely the relation between, on the one hand, the perceived economic and cultural effects of immigration and, on the other hand, attitudes towards immigration policy. Consider the economic impact first. Three variables can be constructed from the ESS that proxy for one’s beliefs about immigrants’ impact on wages, jobs and the fiscal balance. First, the 2002 wave included a question about how much a respondent agreed or disagreed with the statement that ‘average wages and salaries are generally brought down by people coming to live and work’ in the respondent’s country, on a 1–5 scale. We take an answer of 1 (agree strongly) or 2 (agree) as an indicator that one believes
immigrants are bringing natives’ wages down. Second, waves 2002 and 2014 asked for the degree of agreement with the statement that ‘people who come to live here generally take jobs away from workers,’ on a 1–10 scale. We take an answer between 1 and 4 as an indicator that one believes immigrants are hurting natives’ employment. Third, waves 2002 and 2014 also asked the following question: ‘Most people who come to live here work and pay taxes. They also use health and welfare services. On balance, do you think people who come here take out more than they put in or put in more than they take out?’, on a 1–10 scale. An answer between 1 and 4 is again taken as an indicator that one believes the fiscal impact of immigrants is negative. Table 1 shows how these three variables predict a respondent’s statement that the government should allow few or no immigrants in their country, using a linear probability model with country fixed effects and individual covariates. In 2002, believing that immigrants have a negative impact on all of the three economic dimensions was associated with 40 percentage points higher probability of supporting a restrictive immigration policy (last column). Notice that the correlations become larger between 2002 and 2014.19

Consider the cultural impact next. In both 2002 and 2014, respondents were asked to report the degree of agreement with the statement that the ‘country’s cultural life is undermined by immigrants’. One’s concern about the ‘country’s cultural life’ is a particular aspect (not very well defined, admittedly) arguably bearing little

---

19 Estonia and Slovakia are not present in the 2002 wave, and Greece and Italy are not present in the 2014 wave. Results using the subset of countries present both in 2002 and in 2014 are available from the authors, and show that the unbalanced nature of the panel does not explain the larger correlations in 2014.

---

### Table 1. Projection of attitude towards immigration on beliefs about its economic impact

<table>
<thead>
<tr>
<th>Immigrants:</th>
<th>Allow few or no immigrants</th>
</tr>
</thead>
<tbody>
<tr>
<td>– bring wages down</td>
<td>0.180* (0.010)</td>
</tr>
<tr>
<td>– take jobs away</td>
<td>0.205* (0.009)</td>
</tr>
<tr>
<td>– take more than they put in</td>
<td>0.201* (0.012)</td>
</tr>
</tbody>
</table>

| Individual covariates | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Country fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |

Notes: Linear probability models. The dependent variable is a dummy taking value 1 if a respondent states that the government should allow few or no immigrants in the country. Country fixed effects are included, as well as the following individual covariates: gender, age dummies, marital status, number of children, education, employment status, self-reported trust in other people and belonging to a religious denomination. Due to missing data in the ESS, there are no observations from Estonia and Slovakia in the regressions using the 2002 wave, and there are no observations from Greece and Italy in the regressions using the 2014 wave. Robust standard errors in parentheses, clustered at the country level.

*significant at 1% or better.
connection with the labour market or fiscal impact of migrants and pointing instead to cultural externalities. Table 2 adds this variable to the linear projections of Table 1. The cultural variable is more strongly associated with attitudes towards immigration than the labour market or fiscal variables: believing that immigrants undermine country’s cultural life increases the probability of opposing open borders by about 20 percentage points. In this case, too, the correlation becomes larger between 2002 and 2014. This evidence is consistent with the more sophisticated analyses by Dustmann and Preston (2007) and Card et al. (2012). The authors estimate on ESS and British Social Attitudes Survey data, respectively, empirical models identifying the effects of immigration on wages, welfare provision or net fiscal contribution, and cultural identity (or ‘compositional amenities’) and conclude that the perceived labour market and the net fiscal impact of immigrants are of second-order importance in explaining attitudes towards immigration.

A comparison of Table 1 and Table 2 reveals that the cultural variable absorbs a substantial part of the variation in the dependent variable previously attributed to the three economic variables – the coefficients on these three variables in the last column of Table 2 are about 1/4 smaller than in the corresponding column of Table 1. This means that the cultural and the economic variables are positively correlated. Such positive correlation is represented in Figure 2, which aggregates the data at the country level and plots the unconditional fraction believing that the ‘country’s cultural life is undermined by immigrants’ (vertical axis) against the unconditional fraction believing that immigrants bring natives’ wages down, take away their jobs or take out of the fiscal system more than they put in.

<table>
<thead>
<tr>
<th>Immigrants:</th>
<th>Allow few or no immigrants</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>– bring wages down</td>
<td>0.143*</td>
<td>0.107*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.010)</td>
<td></td>
</tr>
<tr>
<td>– take jobs away</td>
<td>0.154* 0.194*</td>
<td>0.122* 0.155*</td>
<td>0.102* 0.102*</td>
</tr>
<tr>
<td></td>
<td>(0.009) (0.010)</td>
<td>(0.010) (0.010)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>– take more than they put in</td>
<td>0.156* 0.165* 0.128* 0.119*</td>
<td>0.119* 0.119*</td>
<td>0.119*</td>
</tr>
<tr>
<td></td>
<td>(0.011) (0.011) (0.011) (0.011)</td>
<td>(0.011) (0.011) (0.011) (0.011)</td>
<td></td>
</tr>
<tr>
<td>– undermine country’s cultural life</td>
<td>0.234* 0.218* 0.264* 0.215* 0.266* 0.190* 0.237* 0.178*</td>
<td>0.178* 0.178*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.018) (0.018) (0.021) (0.017) (0.019) (0.016) (0.019) (0.016)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual covariates</td>
<td>Yes Yes Yes Yes Yes Yes Yes Yes</td>
<td>Yes Yes Yes Yes Yes Yes Yes Yes</td>
<td></td>
</tr>
<tr>
<td>Country fixed effects</td>
<td>Yes Yes Yes Yes Yes Yes Yes Yes</td>
<td>Yes Yes Yes Yes Yes Yes Yes Yes</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Linear probability models. The dependent variable is a dummy taking value 1 if a respondent states that the government should allow few or no immigrants in the country. Country fixed effects are included, as well as the following individual covariates: gender, age dummies, marital status, number of children, education, employment status, self-reported trust in other people and belonging to a religious denomination. Due to missing data in the ESS, there are no observations from Estonia and Slovakia in the regressions using the 2002 wave, and there are no observations from Greece and Italy in the regressions using the 2014 wave. Robust standard errors in parentheses, clustered at the country level.

*significant at 1% or better.
At the same time, the perceived labour market and fiscal effects of immigrants seem to bear no relationship with estimated effects. The actual cross-country economic effects corresponding to the perception evoked in the ESS survey are difficult to identify. Nonetheless, for the labour market effects, Docquier et al. (2014) provide a set of model-based, mutually consistent estimates of the effects of immigration on natives’ wages and employment for a large group of OECD countries. As for the fiscal effects, Liebig and Mo (2013) provide cross-country estimates of the net fiscal contribution of immigrants in 2007–2008, albeit in a static framework (see Section 2.1). The correlation between perceived and estimated effects is illustrated in Figure 3. Although for the labour market effects the correlation has the right sign, in most countries the wage and employment impacts of immigrants are estimated to be positive, contrary to the perception of negative effects by a large share of natives. For the fiscal effects, instead, the correlation (if any) has the wrong sign.

Taking the estimated effects at face value, how can we explain the stark discrepancy between perceptions and reality? One possibility is that people’s perceptions are largely imprecise. Estimates of economic effects certainly are. Another, more interesting possibility suggested by the lack of correspondence between estimates and perceived reality and, at the same time, the positive correlation between perceived economic effects and
concerns about the country’s cultural life (Figure 2) is that the frequently observed opposition to immigration is justified in terms of its perceived economic effects but it is actually motivated by the kind of cultural externalities we have embedded in the model. For instance, a respondent who is intolerant towards immigrants because of cultural or ethnic reasons may have persuaded himself (or report to the interviewer, as a justification) that immigration has adverse economic effects.

To corroborate this presumption, consider one of the implications of the model: absent cultural externalities, if all immigrants are low skilled then their ethnic identity should be irrelevant for natives. Not so in the data: the expressed preference for a restrictive immigration policy varies with respect to the ethnic origin of immigrants. The ESS question about whether the government should allow immigrants to come and live in one’s country is asked separately for ‘immigrants of same race/ethnic group as majority’ and ‘immigrants of different race/ethnic group from majority’. The respective shares of respondents stating that their country should allow ‘few’ or ‘none’ of a specific type of immigrants is displayed in Figure 4 for year 2014. The population-weighted average fractions across the 22 countries are 31.1% for immigrants of the same group as majority, and 42.2% for immigrants of a different group, a substantial difference suggesting that the cultural identity of migrants matters per se.

Figure 3. Actual versus perceived labour market and fiscal effects of immigrants
Notes: The figure relates the perceived labour market and fiscal effects of immigration (horizontal axis) and the actual effects as estimated by Docquier et al. (2014), percentage effects, and Liebig and Mo (2013), thousands of euros. Linear fit superimposed. Sampling weights are applied to ESS data. Ireland is excluded from this figure because it is an outlier.
Differences in natives’ attitudes by education are also informative. In the model, low-skilled natives oppose immigration because of negative economic and cultural externalities. However, the low-skilled favour immigration because a larger low-skilled population has more political power (formally, a larger weight in the government’s objective function). On the contrary, high-skilled natives oppose immigration because of cultural and political externalities (and some negative perceived fiscal externality, possibly) but...
favour it because of their complementarity in the labour market. Presumably, the cultural externality is smaller for the high-skilled because they have limited social contact with immigrants relative to the low-skilled. Therefore, there should be an education gradient in natives’ opposition to open borders.

This is shown in Figure 5, which displays the fraction of natives wanting to allow for few or no immigrants in their country around 2014, by respondents’ educational attainment. The population-weighted average fractions across the 22 countries are 57.8% for natives with less than a high school degree, 52.7% for natives with a high school degree and 31.9% for natives with a college degree or more.

The bottom line of this empirical analysis is that the kind of cultural externalities that play a key role in our model seem to matter in determining natives’ attitudes towards immigration policy, in addition and possibly more than economic externalities: the ethnic identity of potential immigrants is salient (while it should be irrelevant if immigrants were merely low-skilled workers), and the perceived economic effects, while uncorrelated with estimated effects, are correlated with concerns about the country’s cultural life (a particular aspect that should be orthogonal to the economic impact of immigrants).

5. CONCLUDING REMARKS

We have motivated and studied a model of the effects of immigration flows where cultural externalities are a driving force in addition to standard economic externalities. Contrary to the latter, which favour high-skilled natives and so make some voters favourable to more immigration, cultural externalities make both high- and low-skill natives averse to more immigration, although more so the low-skilled. Evidence from the ESS is consistent with this theoretical structure, supporting our claim that one gains deeper insights into immigration policy by considering the nature and consequences of the cultural externalities generated by immigration. We have also briefly discussed three examples of commitment devices – border infrastructure, immigrants’ political rights and fiscal policy – we may expect natives to acquire in response to the resulting time-inconsistency of the immigration policy. Perhaps it is no accident that these three examples correspond to heated political issues in contemporary Europe. Although the model we have studied is simple and lays no claims to being general, we believe it leads to important insights and, most important, it provides a theoretical framework for more comprehensive studies of immigration policy when cultural concerns play a primary role in shaping attitudes towards migrants in receiving countries. This seems to us a key issue in the present economic and political landscape.
Discussion

Vincenzo Denicolò

University of Bologna

This is an insightful paper that discusses how immigration policies may endogenously change over time as a consequence of the fact that today’s immigrants will have a say on tomorrow’s policies. To capture such endogeneity, the authors assume that policy choices maximize the welfare of an ‘average’ agent. The average is taken across the entire population, which includes both natives and immigrants.

Natives are of two types, high-skilled and low-skilled. Immigrants are low-skilled and therefore are a substitute for low-skilled natives. With a well-behaved, constant-returns-to-scale production function and perfectly competitive markets, the wage of high-skilled workers will be an increasing function of the number of immigrants, while that of low-skilled workers a decreasing function. Therefore, high-skilled workers will favour immigration while low-skilled workers will oppose it. (The authors consider also another mechanism that pushes in the same direction, i.e. a congestion effect in the provision of public goods, but that is not really necessary for the model’s results.)

There are two periods. The focus of the analysis is on the fact that first-period immigrants will add to the number of low-skilled workers, implying that in the second period the average agent will be less favourable to immigration. Therefore, in the second period, fewer immigrants will be admitted than in the presence of commitment. Since this distortion is larger, the greater the number of first-period immigrants, in order to reduce the second-period distortion the number of first-period immigrants will also be distorted downward. Therefore, both first-period immigrants and the total number of immigrants will be lower in the absence of commitment than under commitment (Proposition 2).

If commitment is not possible, how could natives attenuate these distortions? One possibility is to run a fiscal deficit, issuing public debt which must be redeemed in period 2. This raises the value of period-2 immigrants, who will pay taxes thus contributing to finance first-period expenditures. As a result, second-period immigration will be closer to the commitment solution. The model here delivers a prediction that is, in principle, testable: anything else equal, countries with higher public debt should be willing to take more immigrants than countries where public debt is low.

However, many readers may find it counterintuitive that natives would like to commit to take more immigrants in the future. If anything, it seems more natural to imagine that natives may want to commit to let fewer immigrants in, in the future. Can the model be reconciled with such presumption? The answer is yes, and this is probably the most important contribution of the paper.
The authors note that in addition to the wage externality discussed above, immigrants may exert a more direct, cultural externality on natives. That is, natives may dislike having more immigrants around. First-period immigrants, on the contrary, may like it. This implies that first-period immigrants will have preferences towards immigration that are closer to those of the high-skilled natives than to those of the low-skilled ones. If this cultural externality is strong enough, the uncommitted second-period immigration rate will be higher than the one arising under commitment. It will still be true that in the absence of commitment, first-period immigration will be reduced so as to attenuate the second-period distortions, but now the overall effect of a lack of commitment on the total number of immigrants may be positive (Proposition 3).

In this case, the first-period natives may want to run a fiscal surplus, rather than a deficit. Alternatively, they might want to make irreversible choices that raise the second-period cost of letting immigrants in. The authors call such choices ‘building walls’, and take Brexit as a possible example of this.

From this synopsis it should be clear that the paper provides a simple, consistent framework that allows analysing important policy issues which are debated heatedly these days. It also bridges the economics literatures on culture and immigration, which had hitherto grown separately.

In general, I like the simplicity and elegance of the paper’s model. However, I think that policy choices could have been modelled in a different and perhaps more standard way. The paper assumes that policies are determined so as to maximize the utility of the average agent. This is equivalent to maximizing a utilitarian social welfare function, but the paper stresses that the analysis is positive rather than normative in nature. Clearly, one can imagine that the objective function is not a proper social welfare function, but rather a reduced form of some political process that somehow aggregates the views of different agents and in doing so gives a positive weight to everyone’s opinion. However, I believe that the paper’s main results could also be obtained in a more standard framework where agents vote and policies reflect the preferences of the median voter.

To smooth out the effects of immigration on policy in a median voter framework, it would be convenient to add some further heterogeneity among agents. The cultural component of the utility function seems an obvious candidate to carry such heterogeneity. Thus, suppose that the cultural component of utility, \( r^\theta (h_l) \), comprises both a deterministic component that depends on the type of the agent, \( h_l \), like in the present model, and also a random component which follows some probability distribution. Under some regularity conditions, each agent’s preferences over immigration will be single-peaked and hence the median voter theorem will apply.

In this framework, immigration changes the median voter, and hence the equilibrium policy. Qualitatively, it does so in the same way as in the present model. Thus, if the positive cultural externality that new immigrants exert on old ones is negligible, we are back to the baseline model where more immigration today causes less immigration tomorrow. If instead the cultural externality is sufficiently strong, this result may be reversed. The results would therefore be the same as in the current model. The choice of a utilitarian
social welfare function or the median voter is then largely a matter of taste, and I personally believe that the median voter formulation would be cleaner.

Mikhail Drugov
New Economic School

The main point of this paper is simple but important. Immigrants accepted today will have some influence in the political process in the future and in particular, in the decisions concerning future immigration policy. Hence, deciding on the current immigration policy the natives take into account the fact that the future immigration policy will be biased towards the preferences of today’s immigrants. This general point is well known in the literature; Acemoglu et al. (2012) summarize it very succinctly: ‘At the center of our approach is the natural lack of commitment in dynamic decision-making problems - those that gain additional decision-making power . . . cannot commit to refraining from further choices that would hurt the initial set of decision makers.’ (p. 1447).

Bizin and Zanella’s main results characterize the distortions that this dynamic inconsistency will bring. It may seem intuitive that immigrants would press for accepting more immigrants in the future; anticipating this, the natives reduce the inflow of immigrants today. This is indeed what happens when the cultural externalities are important. It is assumed that immigrants exert negative cultural externalities on natives and positive ones on other immigrants. However, (almost) the opposite happens when economic externalities are crucial. When the migration is mostly low-skilled, high-skilled natives benefit economically from it while the low-skilled natives lose. Hence, high-skilled natives like migration and the low-skilled ones oppose it. Immigrants then have the same preferences as low-skilled natives and the future immigration policy will be biased towards less migration. Interestingly, however, the first-period migration is still lower than in the ‘commitment’ policy which is obtained when immigrants do not get any influence in the second period.

This is an important theoretical point and it is interesting to think what it means in practice. The paper provides evidence from the ESS consistent with the cultural externalities (the natives’ attitudes depend on the immigrants’ origin) and showing that respondents with higher educational attainment are more favourable to immigration. This is the evidence about the assumptions of the model and it is fairly uncontroversial. But there is no discussion of whether or how this time inconsistency, that is, the difference between the equilibrium and the commitment policies can be detected and measured. The major issue is that the commitment policy is not observed in reality. In other words, we need to measure the difference between the actual policy and some other unobserved one which does not seem easy. One way could be as follows: as discussed in the paper, the natives might look for commitment devices, that is, for various ways to constrain the immigrants’ future influence by affecting future costs and benefits of various options or by directly limiting their political rights. The paper mentions several such
devices such as a direct decrease in the marginal cost of containing future migration (‘building walls’?) and a tight fiscal policy. A larger difference between the equilibrium and the commitment policies creates a higher demand for some commitment device (the paper does not really provide comparative statics results on the size of the time inconsistency but it is easy to obtain at least in some cases) though there should be also an argument why, say, a ‘wall’ would not be built under the (unobserved) commitment policy.

A related issue is that the direction of the time inconsistency is ambiguous as discussed above. Hence, it is possible that an increase in cultural externalities brought about by, say, the changing origin of the migrants or by migration becoming more salient, counteracts the economic externalities making the total migration under the equilibrium policy closer to the one under the commitment policy.

As a result, it is not immediately clear what policy-makers should do once they read this paper and realize the potential time inconsistency of the immigration policy. A further empirical framework is definitely needed. Also, immigrants will affect any future policy and there might be important effects concerning other policies too, such as redistribution or the role of religion in the society.

Panel discussion

John Kennan wondered whether it is useful to use social welfare when describing the model. He argued that what the authors are investigating is a decision-making process. As a result, the social welfare function introduces a lot of ambiguity in terms of who counts, e.g. what is the welfare of people that is not admitted into a country? In response to John Kennan’s comment, Giulio Zanella acknowledged that the social welfare terminology may indeed be misleading.

Kevin O’Rourke observed that the late nineteenth-century evidence regarding voting on immigration restrictions in the United States finds that legislators were more likely to vote for more liberal immigration policies in cities with more immigrants. Camille Landais said it would be important to have a sense of the magnitude of the externalities the paper refers to since one would need very large externalities for this to become a first-order issue.

Christian Dustmann noted that a very important assumption in the model is that preferences of immigrants with respect to cultural capital are different from those of migrants. However, evidence suggests that immigrants quickly adapt in terms of preferences to that of the receiving country. Giulio Zanella clarified that there are many ways to generate time inconsistency but recognized that this assumption needs to be better justified in the paper.
APPENDIX: PROOFS OR PROPOSITIONS

Proposition 1. The social welfare choice problem is generically time-inconsistent; that is, the no-commitment solution for immigration flows, \{m_{i1}^{nc}, m_{i2}^{nc}\}, is generically distinct from the commitment solution, \{m_{i1}^{c}, m_{i2}^{c}\}.

Proof. Time-inconsistency follows directly from the comparison of conditions (9) and (11), respectively, for the commitment and the no-commitment cases. Condition (9) is satisfied at the unique commitment solution, while any solution of the no-commitment problem instead requires (11). By the Inada conditions, (9) is satisfied at the unique commitment solution, while any solution of the no-commitment case satisfies (11). By the Inada conditions, we can without loss of generality restrict our analysis to \(m_1 > 0\) in both the commitment and the no-commitment case. Note that \(r^0(p) = 0\), \(\theta = h, l, i\), implies that the commitment problem depends only on \(q_l\) (the stock of immigrants) not on \(p_l\) (the stock of non-assimilated immigrants). As a consequence, \(V_2(q, p) = V_2(q)\) and \(\frac{\partial V_2(q)}{\partial m_1} = \frac{\partial V_2(q)}{\partial m_2}\). Replacing into (8) and (9), implies then that the unique solution to the commitment problem satisfies \(\frac{\partial V_i(m_1)}{\partial m_1} = 0\). Remember that the weights are normalized so to sum-up to 1 and that immigrants have zero weight in the first period:

\[
m_{i1}^{nc} + m_{i2}^{nc} < m_{i1}^{c} + m_{i2}^{c}.
\]

Moreover, the immigration flow is smaller at equilibrium (with no-commitment) than with commitment:

\[
m_{i1}^{nc} < m_{i1}^{c}.
\]

Proof. Under the Inada conditions, we can without loss of generality restrict our analysis to \(m_1 > 0\) in both the commitment and the no-commitment case. Note that \(r^0(p) = 0\), \(\theta = h, l, i\), implies that the commitment problem depends only on \(q_l\) (the stock of immigrants) not on \(p_l\) (the stock of non-assimilated immigrants). As a consequence, \(V_2(q, p) = V_2(q)\) and \(\frac{\partial V_2(q)}{\partial m_1} = \frac{\partial V_2(q)}{\partial m_2}\). Replacing into (8) and (9), implies then that the unique solution to the commitment problem satisfies \(\frac{\partial V_i(m_1)}{\partial m_1} = 0\). Remember that the weights are normalized so to sum-up to 1 and that immigrants have zero weight in the first period:

\[
\psi_{i1}^{k} + \psi_{i1}^{l} = 1,
\]

\[
\psi_{i2}^{k} + \psi_{i2}^{l} + \psi_{i2}^{l} = 1.
\]

As a consequence, \(m_1 > 0\) implies \(\psi_{i2}^{k} < \psi_{i1}^{k}, \psi_{i2}^{l} < \psi_{i1}^{l}, \psi_{i2}^{l} > \psi_{i1}^{l}, \text{and } \psi_{i2}^{k} + \psi_{i2}^{l} > \psi_{i1}^{k}\). Given weights \(\psi_{i2}^{k}, \) any solution for the no-commitment problem satisfies Equation (11). Also, for given weights \(\psi_{i2}^{k}, \) the no-commitment problem, too, depends only on \(q_l\) and not on \(p_l\). Recall that \(\{m_{i1}^{c}, m_{i2}^{c}\}\) satisfies \(\frac{\partial V_2(q)}{\partial m_2} = \frac{\partial H_2(q)}{\partial m_2}\), while \(m_{i2}^{nc}\) satisfies \(\frac{\partial V_2(q)}{\partial m_2} = \frac{\partial H_2(q)}{\partial m_2}\). But, in this environment, \(V_2(q_2) = \sum_{\theta=h,l,i} \psi_{i1}^{\theta} v_{i1}^{\theta}(q_2), \) and \(H_2(q_2) = \sum_{\theta=h,l,i} \psi_{i2}^{\theta} v_{i2}^{\theta}(q_2), \) where \(q_2 = m_1 + m_2.\) Furthermore, under our assumptions, \(v_{i1}^{l}(q_2) = v_{i2}^{l}(q_2).\) Using these
facts, compare first order conditions (9) and (11), which are rewritten here explicitly for easier reference:

\[
\psi_1 \frac{\partial \psi_2(m_1 + m_2)}{\partial m_2} + \psi_1' \frac{\partial \psi_2(m_1 + m_2)}{\partial m_2} = \frac{\partial \psi(m_1 + m_2)}{\partial m_2}, \tag{9}
\]

\[
\psi_2 \frac{\partial \psi_2(m_1 + m_2)}{\partial m_2} + (\psi_2' + \psi_2^i) \frac{\partial \psi_2(m_1 + m_2)}{\partial m_2} = \frac{\partial \psi(m_1 + m_2)}{\partial m_2}. \tag{11}
\]

The change in weights implies that, for a given \( q_2 \), the LHS of (11) decreases relative to the LHS of (9). This is so because under our assumptions \( \frac{d\psi_2(q_2)}{dm_2} < \frac{d\psi_1(q_2)}{dm_2} \), for any \( m_2 \). Therefore, the concavity of \( \psi_2(.) \) and the convexity of \( \psi_1(.) \) imply that \( m_1^c + m_2^c \) satisfies (11) if \( m_1^c + m_2^c < m_1^i + m_2^i \). This result holds for any weight structure \( \psi_2^h < \psi_1^h, \psi_2^i < \psi_1^i, \psi_2^i > \psi_1^i \), that is, for any weight structure associated with \( m_1^i > 0 \). This proves the first part of the proposition.

To prove the second part, notice that Equation (10) would be satisfied for a \( m_1^c = m_1^i \) if the social welfare weights were constant in \( m_1 \); in this case \( \frac{dm_2}{dm_1} = -1 \) (because in this case \( m_1 + m_2 \) would be the same under commitment and under no-commitment) and, as a consequence, the no-commitment problem, too, would satisfy \( \frac{\partial \psi_1(m_1)}{dm_1} = 0 \). However, the social welfare weights are not constant in \( m_1 \). Accounting for the dependence of the social welfare weights on \( m_1^i > 0 \), it must be \( \frac{dm_2}{dm_1} < -1 \) (because in this case \( m_1 + m_2 \) is smaller under no-commitment than under commitment, as proved above). Equation (10) is then satisfied for \( m_1^c < m_1^i \).

**Proposition 3.** Allowing for cultural externality \( r^\theta(p), \theta = h, l, i \), but with no-assimilation, \( p_t = q_0 \) the total immigration flow is larger at equilibrium (with no-commitment) than with commitment if the cultural externality is sufficiently large for past immigrants:

\[
m_1^c + m_2^c > m_1^i + m_2^i \text{ if } \frac{dr^\theta(q)}{dm_2} \text{ is sufficiently large}
\]

However, the immigration flow is smaller at equilibrium (with no-commitment) than with commitment in the first period:

\[
m_1^c < m_1^i.
\]

**Proof.** The proof is analogous to the proof of Proposition 2. In this environment it is still the case that the commitment problem depends only on \( q_b \), and so the unique solution to the commitment problem satisfies \( \frac{\partial V_1(m_1)}{dm_1} = 0 \). Under the Inada conditions, we can still restrict to \( m_1 > 0 \) in both the commitment and the no-commitment case and as a consequence, \( \psi_2^h < \psi_1^h, \psi_2^i < \psi_1^i, \psi_2^i > \psi_1^i \), and \( \psi_2^i + \psi_2^i > \psi_1^i \). Furthermore, it is still the case that, for given weights \( \psi_2^0 \), the no-commitment
problem, too, depends only on \( q_c \). In this environment, however, differently from the case with no cultural externality, \( V_2(q_2) = \sum_{\theta=h,l,i} \psi_1^d q_2 + r^d(q_2) \) and \( H_2(q_2, p_2) = \sum_{\theta=h,l,i} \psi_2^d q_2 + r^d(q_2) \). Under our assumptions, \( \psi_2^l(q_2) = \psi_2^d(q_2) \) and \( \frac{d[\psi_2^l(q_2)+r^l(q_2)]}{dm_2} < \frac{d[\psi_2^d(q_2)+r^d(q_2)]}{dm_2} \), for any \( m_2 \). Compare first-order conditions (9) and (11), reproduced here explicitly like in the proof of Proposition 2:

\[
\begin{align*}
\psi_1 \frac{\partial [v_2^q(m_1 + m_2) + r_2^d(m_1 + m_2)]}{\partial m_2} + \psi_1 \frac{\partial [v_2^q(m_1 + m_2) + r_2^d(m_1 + m_2)]}{\partial m_2} &= \frac{\partial x(m_1 + m_2)}{\partial m_2}, \\
\psi_2 \frac{\partial [v_2^q(m_1 + m_2) + r_2^d(m_1 + m_2)]}{\partial m_2} + \psi_2 \frac{\partial [v_2^q(m_1 + m_2) + r_2^d(m_1 + m_2)]}{\partial m_2} &= \frac{\partial x(m_1 + m_2)}{\partial m_2} \\
-\psi_2 \frac{\partial [v_2^q(m_1 + m_2) + r_2^d(m_1 + m_2)]}{\partial m_2}.
\end{align*}
\]

Extending the reasoning behind the proof of Proposition 2, these imply that if \( \frac{d[\psi_2^l(q_2)+r^l(q_2)]}{dm_2} \) is sufficiently larger than \( \frac{d[\psi_2^d(q_2)+r^d(q_2)]}{dm_2} \), that is, if \( \frac{dv(q_2)}{dm_2} \) is large enough, then \( m_1^{mc} + m_2^{mc} > m_1^c + m_2^c \). Otherwise, \( m_1^{mc} + m_2^{mc} < m_1^c + m_2^c \). In this environment, therefore, the weight structure might affect whether the \( m_1 + m_2 \) which solves (11) is greater or smaller than \( m_1^c + m_2^c \). This proves the first part of the proposition.

To prove the second part, suppose that at the weight structure determined by \( m_1^c \) it is \( m_1^{mc} + m_2^{mc} > m_1^c + m_2^c \). Equation (10) is then satisfied for a \( m_1^{mc} = m_1^c \) if \( \frac{dm_2}{dm_1} = -1 \). This would be the case if social weights were constant in \( m_1 \). But accounting for the dependence of social welfare weights on \( m_1 > 0 \), it is \( \frac{dm_2}{dm_1} > -1 \) (because in this case \( m_1 + m_2 \) is larger under no-commitment than under commitment). Equation (10) is then satisfied for a \( m_1^{mc} < m_1^c \). If instead, at the weight structure determined by \( m_1^c \), it is \( m_1 + m_2 < m_1^c + m_2^c \), with \( \frac{dm_2}{dm_1} < -1 \), then Equation (10) is still satisfied for a \( m_1^{mc} < m_1^c \).

\[\square\]

REFERENCES


