Full title: Which transition comes first? Urban and demographic transitions in comparative perspective

Short title: Urban and demographic transitions

Authors and affiliation:

Philippe Bocquier, Centre de recherche en démographie et sociétés, Université Catholique de Louvain, Belgium

Rafael Costa, Centre de recherche en démographie et sociétés, Université Catholique de Louvain, Belgium

Address for correspondence:

Prof. Philippe Bocquier
Université Catholique de Louvain
SSH/IACS, Collège Jacques Leclercq
Place Montesquieu 1, bte L2.08.05
1348 Louvain-la-Neuve
Belgium
philippe.bocquier@uclouvain.be
Tel: +32 10 47 41 83
Fax: +32 10 47 29 52

Abstract

This paper is an attempt to shed light on the relationship between demographic and urban transitions in 19th century Europe, taking Sweden and Belgium as examples. Results reinstate migration as the major component of early urban transition. Theoretical consequences are then drawn on the role of spatial economic differentials in demographic transitions.

Keywords: Demographic transition; urban transition; migration; economic development.
Introduction

Analysing demographic transition is a recurrent task in demography. Although one might regret that demographic transition is not sufficiently scrutinised, and that no sufficient historical data are available to that purpose, the Population and Development Review 2011 special issue on demographic transition is a commendable initiative to re-examine demographic transition theory and confront it to empirical evidences (Lee and Reher 2011). The present paper is a follow-up to this initiative and in particular to the attempt of Tim Dyson in the same PDR special issue to link demographic transition and urbanization in a comprehensive theory (Dyson 2011). The first section of the present paper revisits the theory of demographic transition in the light of Zelinsky (1971), de Vries (1990), Chase-Dunn (1998) and Galor (2011). The second section examines the components of urban-rural growth differences and offers an alternative to Dyson’s analysis on Sweden and is enriched using new data on Belgium using the same methodology. Our analysis reconsiders the role of migration in urban transition. Theoretical considerations on the relation between these transitions and economic development are then drawn.

Demographic transition and urban development

Very few authors before W. Zelinsky, a trained geographer, have theorized the spatial dimension of the demographic transition. He established for the first time the importance of spatial analysis in the study of the demographic transition in his seminal work (Zelinsky 1971: in this section all page numbers are in reference to his article, unless otherwise stated). Zelinsky considered demographic transition as a diffusion process that is at the same time temporal and spatial. Several analogue terms such as ‘pandemic’ or ‘tumorous cells’ are used to explain the propagation of new demographic behaviours across space. The author justifies his spatial

\[\text{\footnotesize The terminology of the 1960s used by Zelinsky needs sometimes to be translated into our contemporary language (e.g. developed countries for “advanced societies”).}\]
Urban and demographic transitions - 2

approach by the conviction that spatial distribution has a specific meaning (geographical axiom), by the concept of spatial diffusion of human innovations (anthropological axiom), and by the principle of least effort (economic axiom) (pp. 219-221). His article posits that “there are definite, patterned regularities in the growth of personal mobility through space-time during recent history [Zelinsky refers implicitly to the twentieth century], and these regularities comprise an essential component of the modernization process” (pp. 221-222).

The insistence on the role of spatial mobility, which was new at the time, and the use of the term ‘mobility transition’ in the title is at the origin of a frequent misunderstanding of Zelinsky’s work. Most commentators assumed that his paper was about migration and often missed his main point that mobility is embedded in, and not just parallel to, demographic transition. Zelinsky actually prefers ‘vital transition’ to name the conventional perspective that considers only fertility and mortality as the components of demographic transition. Demographic transition à la Zelinsky is actually the integration of ‘vital transition’ and ‘mobility transition’, in an all-inclusive acceptance. It is actually abusing the term ‘demographic transition’ to refer to fertility and mortality transitions only. As D. Courgeau and R. Franck remind us (2007) demography is the study of the combinations of fertility, mortality and migration that explain population growth, decline or stabilization. Many other individual and social level factors can influence population variations but the analysis of demographic parameters (namely fertility, mortality and migration) is a perspective particular to demography, say Courgeau and Franck (2007, pp. 39-40). Zelinsky’s terminology will be used in the present text, referring to ‘vital transition’ to mean interplay of crude birth and death rates, and to ‘mobility transition’ to mean the increase in population mobility, including migration. The term ‘demographic transition’ is therefore the sum of the ‘vital transition’ and ‘mobility transition’. The term ‘urban transition’ refers to the change from a mainly rural to a mainly urban society.
Zelinsky actually establishes a correspondence between the different phases of the vital and mobility transitions (Table 1, p230). He mentions that this correspondence had been tested by Friedlander (1969), who wanted to verify the hypothesis by Davis (1963) according to which fertility levels are inversely proportional to migration opportunities. Zelinsky extends this relational hypothesis to the three demographic parameters although he acknowledges that the empirical proof may reveal notable differences in time scale and intensity depending on the geographical and historical context. The correspondence established by Zelinsky between the vital and the mobility transitions indicates their contemporaneousness and probably their interdependence (p.229). Several graphs (p.228) express the complementary nature and the intertwinement of the temporal and spatial dimensions of vital and mobility transitions, unified in the demographic transition. As these transitions are continuous phenomena by construction – sudden breaks are the exception – the limits between the different phases are partly arbitrary and are meant only to ease presentation. The correspondence that Zelinsky establishes between transitions is only indicative of an ‘ideal nation’ in his own words (echoing the ‘ideal type’ à la Max Weber), a historical mean of developed countries. Zelinsky recommends empirical studies to document the transition periods and the different mobility patterns that are sketched in his original article (p.233).

Although Zelinsky is rather vague on the precise mechanisms that link vital and mobility transitions, he posits that demographic transition (i.e. vital and migration transition) and other transitions (social and economic) are mutually interdependent and calls for a “general-systems approach” (p.249). There would be no direct causal relationship between these transitions but, as we would say today, a systemic relationship. In Zelinsky’s explanation of urbanization, “the implicit assumptions seems to be that the modernization driving the demographic transition also loosens the constraints that immobilize labour, unleashing rural-to-urban migration for reasons unrelated to urban labour markets” (de Vries 1990, p.54). Urbanization appears to be explained mainly by mobility transition while the contribution of vital transition is ignored.
Two flaws have been identified in Zelinsky’s theory. First, the evolutionist perspective that considers past societies as immobile has been questioned (Skeldon 1997). Historical inaccuracies, especially regarding the assumed low and inconsequential mobility of ancient or developing societies, have been identified (Lucassen and Lucassen 2009). Second, urbanization may result from the vital transition, i.e. from a difference of natural increase between urban and rural areas (de Vries 1990). In other words, migration contribution to urbanization might neither have been so crucial nor so new, while the role of natural increase has been neglected.

In Jan de Vries’ theory of urban transition (de Vries 1990) urban mortality plays a key role in European countries. The stylized urban transition model is (Dyson 2011): A ‘demographic sink’, when deaths exceed births was observed in urban areas in pre-transition time (this corresponds to the urban penalty characterized by high mortality identified by Gould 1998); Reduction of infectious diseases in urban areas triggered urban natural growth, which itself allowed urban growth; Rural-urban migration was present before urban growth and really played a role in urban growth only after urban mortality declined; This role was further enhanced when mortality declined in rural areas.

Fox (2012) points out that innovation (in particular in agricultural productivity, transportation, and health) should not be confused with economic growth (as measured by GDP per capita) when analysing the forces behind urban transition. In presence of innovations, there may be urban growth without economic growth as was observed in sub-Saharan Africa or other developing areas of the world. Population growth may have a higher contribution than economic growth on urban transition. In particular, mortality decline may be a sufficient condition for urban growth, provided that innovation is sufficient to create a surplus of energy (food and fuel). Fox’s analysis of the effects of geographical characteristics on proportion urban and urban population size confirms that energy surplus occurs more often in coastal areas and river valleys where this surplus is easier to achieve (Fox 2012). One of his conclusions is: “The onset of the
urban transition in any given country or region should therefore be understood as part of a global historical process linked to technological and institutional change and diffusion, not simply as a product of endogenous economic and demographic forces.” (p293). However, we want to argue here against his next conclusion: “The fact that many of the technological and institutional changes that drive mortality decline and facilitate surplus expansion also drive economic development is the source of the spurious conclusion that urbanization is fundamentally a byproduct of economic development.” (p294).

Economic development is conceived in Fox’s and other authors’ papers as national income growth. First, economic development cannot be reduced to income growth. Structural economic changes – including industrialisation and, later, tertiarisation – may occur without income growth. But more importantly, although they are usually measured at country level, these changes originate in international processes. A good example of this process provided by Fox (2012) himself is colonial investments which he identifies as strong drivers of 1950s urbanisation in Africa (p296-297). Starting with colonisation, economic development is more and more globalised so that comparing a 19th century European country with a 21st century African country is meaningless unless one accounts for the very different global environment in which these countries evolved. Economic, demographic, and urban development in a 19th century European country may be to some extent considered independent from that of countries in other continent of the world. However, development in today’s African countries cannot be considered independent from global development. Considering each and every country as separate entities across space and time is subjecting the analysis to “methodological nationalism” bias. National indicators may be relevant for 19th century European countries to explain various aspects of development (although European countries’ development were not independent from each other), but are certainly not for 20th or 21st centuries countries. Other dimensions, such as global economic dependency and country-to-country growth differentials, would need to be included in the analysis of each country’s development. Urbanisation in an average African country may not
be the product of industrialisation *in this country* but may serve (and therefore originate) in the industrialisation of *other countries* elsewhere in the world, Europe in the 19th or 20th centuries, the rest of the world in the 20th or 21st centuries. In other words, development whatever its dimension is less and less an endogenous (national) process and more and more an interrelated (global) process. As time goes by, the scale has to be changed from national to international.

In this paper, we will not make attempt at analysing 20th century urbanisation in the world in relation to their alleged determinants. We will rather look at the particular sequence of demographic and urban transition in 19th century Europe. We will focus on the contribution of the vital and the mobility transitions to the urban transition. The vital transition (i.e. the interplay of fertility and mortality transitions) will be taken as a proxy for disease control and energy surplus, while the mobility transition (i.e. net migration) will be taken as a proxy for economic development. The sequence of the transitions are analysed to show which transition is a necessary (but not necessarily sufficient) condition for the next transition to occur. Note that a sequence is indicative of a possible causal relationship between transitions but not a proof of this relationship, since a third cause may explain both transitions. Following Fox’s theoretical framework of the urbanisation process, this underlying, often unobserved, cause is brought about by technological and institutional innovations. Although descriptive, our analysis will show that mobility transition (i.e. structural economic change) more than demographic transition is at the origin of urban transition.

**Verifying hypotheses on urban and vital transitions in Sweden and Belgium**

Dyson’s analysis of historical urban and rural demographic series on Sweden (1750–1956) and Sri Lanka (1891–1964) seem to conform to the hypothesis that mortality decline drove urban transition. Dyson used series of crude birth and death rates by urban and rural residence to
conclude that in Sweden “until about 1850 the urban death rate was higher than the urban birth rate [...] therefore the urban rate of natural increase was negative. [...] So it was rural natural increase, through migration, that maintained [...] limited urban growth [...] Lower urban fertility also helped to explain why the urban sector was a sink. The urban death rate [fell from about 1800] much faster than the rural death rate” (Dyson 2011, p.43, Figure 2).

Barring some data quality issue, these two analyses seem to confirm that “mortality decline is the remote (i.e. underlying) cause of fertility decline” (Dyson 2011, p.37). However, they are not sufficient to confirm a hypothesis that both de Vries and Dyson make about the relation between urban transition and demographic transition: vital transition and not economic growth explain urban transition since “in recent decades urbanization has been happening in places where there is little or no economic growth—in particular, sub-Saharan Africa” (Dyson 2011, p.38). “Urbanization is both an integral component and an outcome of the demographic transition” (Dyson 2011:47). The fact that “the speed of urbanization [in sub-Saharan Africa and Latin America] seems to have been significantly faster than was the case in developed countries historically [appears to confirm the expectation that a country would] urbanize faster under conditions of faster population growth” (Dyson 2011, pp. 48-50, Figure 4).

Our contention is that time series analysis of crude birth and death rates by sector presented by Dyson are not appropriate to confirm his hypothesis for two reasons:

- These series do not account for the contribution of migration in urban transition. Because rural-to-urban migration is linked to economic growth in the urban sector, migration is a crucial indicator that needs to be estimated to evaluate the role of the economy in the urban and demographic transitions.

- These series do not show the difference between urban and rural growths resulting from the vital and urban transitions. Without these figures and their relations to migration, the role of the demographic transition in the urban transition cannot be fully apprehended.
Although T. Dyson’s analysis of historical demographic series on Sweden and on Sri Lanka seem to conform his hypothesis, we will show how a re-examination of the basic demographic equation may lead to more nuanced conclusions based on the same data.

In the basic (or balancing) demographic equation:

\[ P_{t+h} - P_t = B_h - D_h + I_h - E_h \]

the population \( P \) is circumscribed to a space and the outside world is defined by default (non-
\( P \)) as the population from which immigration is originated or to which emigration is directed.

The non-\( P \) population is ignored in this equation. Now consider two related spaces, say \( u \) for urban and \( r \) for rural, in a modified set of basic demographic equations, assuming same time interval \( h \) for simplicity:

\[
\begin{align*}
P^r_{t+h} - P^r_t &= B^r_h - D^r_h + M^u_r - M^{ur}_h + M^r_h - M^r_r \\
P^u_{t+h} - P^u_t &= B^u_h - D^u_h + M^r_u - M^{ur}_h + M^u_h - M^u_r
\end{align*}
\]

where \( M \) stands for migration flow from \( u \) to \( r \) and from \( r \) to \( u \), or from non-\( P \) (noted ‘\( r \)’) to \( r \) or \( u \), and from \( r \) or \( u \) to non-\( P \). This set of equations makes it possible to analyse the demographic transition in a heterogeneous population \( P = P^u + P^r \). Rates of increase in time interval \( h \) in rural and urban areas are defined as:

\[
\begin{align*}
q^u_h &= \frac{B^u_h - D^u_h + M^{ur}_h - M^r_u + M^u_h - M^{ru}_h}{P^u_t} = b^u_h - d^u_h + m^u_h + i^u_h \\
q^r_h &= \frac{B^r_h - D^r_h + M^{ru}_h - M^{ur}_h + M^r_h - M^r_r}{P^r_t} = b^r_h - d^r_h + m^r_h + i^r_h
\end{align*}
\]

---

\(^2\) We are grateful to Tim Dyson who has kindly provided us with the data that he used in his PDR 2011 article. That way, the differences in our analyses are not flawed by differences in data sources. The details on sources and their limitations can be found in Dyson’s article.
where the denominator \( P_t \) is the population at the beginning of the period but can indifferently be replaced by the mid-interval population. Prospective rates are just easier to use as probabilities for projection purposes. The rate \( b_h, d_h, m_h, i_h \), are respectively the crude birth rate, crude death rate, rate of internal migration increase, and rate of international migration increase, all in period \( h \). Comparing trends of mortality, fertility and migration rates from both areas are of particular interest for the joint analysis of the demographic and urban transitions. From previous equations we derive the urban-rural difference in population growth as:

\[
q_h^u - q_h^r = (b_h^u - b_h^r) - (d_h^u - d_h^r) + (m_h^u - m_h^r) + (i_h^u - i_h^r)
\]

To ease reading of the Figures below, each difference in urban-rural rates can be abbreviated:

\[
URGD = URCBRD - URCDRD + URMGD + URIGD
\]

where URGD stands for urban-rural growth difference, URCBRD for urban-rural growth difference in crude birth rate, URCDRD for urban-rural growth difference in crude death rate, URMGD for urban-rural internal net-migration growth difference, and URIGD for urban-rural international net-migration growth difference. URMGD and URIGD represent respectively the internal migratory and the international migratory attraction or repulsion of urban areas. One may be tempted to consider the difference between urban and rural international net-migration negligible. However this hypothesis of negligible demographic interaction with the outside world is not necessary, since the sum of URMGD and URIGD represents the general migratory attraction or repulsion of urban areas, whatever the origin of migration flows. We will show the importance of considering the non-\( P \) population when analysing demographic transition in developing countries later in the paper.

Interestingly, from a computational point of view, it is sufficient to know the urban-rural growth difference (URGD) and the urban-rural natural growth difference (URNGD=URCBRG-URCDRG) to deduct the urban-rural net-migration growth difference (URMGD+URIGD),
without knowing the exact migration flows between rural and urban areas, and between these areas and the outside world, flows that are admittedly hard to collect:

\[ URMGD + URIGD = URGD − (URCBRD − URCDRD) \]

Using this interesting feature of the two-sector demographic equation, we graphed each parameter using Dyson’s data on Sweden (Figure 1) and using our own data on Belgium (Figure 2), and we added the urban-rural natural growth difference (URNGD). Plotting urban-rural differences in vital and migration rates allow a much easier identification of the components (birth, death and migration) of urban-rural growth difference and therefore of the urban transition that occurs when urban areas grow faster than rural areas. To note, we are here not so much interested by the components of urban growth per se than by the components of urban-rural growth difference.

[Figure 1 & Figure 2 around here]

Data quality issues are already commented at length in Dyson’s paper. We will consider his data as good enough to serve his purpose and ours. Although no precise series before 1820 are available to confirm urban trends in Sweden, we will consider as Bairoch (1988) and Dyson himself (2011) that the percentage urban was constant between 1750 and 1820. Bairoch estimated it at 6.6%, but we will use Dyson’s estimation (9.8%) for consistency with his own demographic series on Sweden.

Data on Belgium were never published before. Since the birth of the Kingdom of Belgium in 1830, information on births, deaths and, to a certain extent, migrations have been collected through vital registration and registers of communes. The Central Commission for Statistics has been collating this information in statistics books since 1941, available to us in different forms: original paper copies (1841 to 1848), microfilms (1857 to 1879) and electronic files (1880 to 1996). To note, the administration did not collate statistics during the 1849-1856 period. There were about 2500 communes from 1841 to 1964 after which old communes were merged into 589
new communes. The average communal area before 1965 was 11.7 km², with 14.6% more than 20 km² and only 2.5% more than 40 km². The official definition of urban areas used from 1867 is applied from 1841 to 1964: all communes with more 5000 inhabitants are considered urban.

Before 1867, the urban communes had to be reconstructed from the original data to comply with the 5000 threshold. This tedious work has been done for a sample of years: 1841, 1844, 1846, 1848, 1857, 1861 et 1864. From 1867, series are available for each year but are not reliable in the world war periods and the years immediately after the First World War. The years 1914-1923 and 1940-1944 are therefore excluded from the analysis. Population estimates are often biased due to the unreliable registration of migrations. In particular out-migrations are badly registered, leading to an overestimation of the population. However, this was corrected using census data by redistributing the difference between the registers’ and censuses’ estimates over the inter-census period. Censuses were conducted every ten years from 1846 to 1876 and from 1880 to 1930, and in 1947, 1961, 1970, 1981, 1991, and 2001.

Figure 1 and Figure 2 not only depict the difference in vital transition between urban and rural areas, but also the parallel trends in urban-rural net-migration growth difference (URMGD+URIGD). For Sweden, Figure 1 shows that positive migration growth was very likely compensated by negative natural growth, resulting in virtually zero urban growth from 1750 to 1836. The ‘demographic sink’ (when URNGD was at its lowest) reached its deeper point (-1.71% difference annual natural growth at the expense of urban areas) around 1815. It is assumed to have been compensated by rural-to-urban migration, leading to a zero urban growth. Confirming Dyson (2010), higher crude death rate in urban areas compared to rural areas explains almost entirely this urban demographic sink. Urban-rural differences in crude birth rate are small over the whole study period. From about 1840 on, the urban-rural natural growth difference reduced and from 1886 on hovered around zero (plus or minus 0.3%). But the most important feature of the Figure 1 is that migration contributed to the urban-rural total growth difference (URGD) and its variations in an essential manner. Certainly urban natural growth became positive thanks to
the reduction of urban-rural mortality differential in the early nineteenth century, but urban natural increase did not supersede rural natural increase in a very significant way. Contrary to de Vries hypothesis, the rising urban natural growth did not ‘diminish the importance of rural-urban migration as a source of urban growth’ (de Vries 1990, p.58). Rather, the quasi-balance of urban and rural natural growth released the full potential of an already existing migration impact on total urban growth. What the Swedish example teaches us is that migration not only intervened in an early phase of urban development before the nineteenth century (proto-industrialization), but was an essential component of urban growth throughout the nineteenth century. If we may use a mechanic metaphor, the engine of the urban transition in Sweden was already running high with migration fuel when the natural growth brakes were released around 1840 thanks to a reduction in urban-rural mortality differential, allowing urban transition to run at full speed.

Trends in Belgium (Figure 2) contrast interestingly with those of Sweden. In both countries, the crude death rate was higher in urban areas before 1910, but the difference reduced almost linearly from about 1880. However, unlike Sweden, the crude birth rate was higher in urban areas before 1900 in Belgium, leading to almost negligible urban-rural differences in natural growth. In other words, higher fertility compensated for higher mortality: there was no pre-transitional ‘demographic sink’ in Belgium. It seems that the vital transition occurred essentially in the 1880-1900 period and more rapidly in urban areas than in rural areas. From 1910 on, differences in crude death rates were negligible while crude birth rates were constantly lower in urban areas: there was actually a post-transitional ‘demographic sink’ in Belgium, but due to fertility not mortality. But the most noticeable feature of trends in Belgium is the role of migration in overall urban growth. First, as in Sweden, migration was an essential component of urban growth throughout the nineteenth century. Migration contributes almost entirely to urban growth over the study period (migration and total differential curves are almost undistinguishable). Second, contrary to Sweden, the vital transition has very little impact on urban growth: the absence of
‘demographic sink’ due to mortality allowed the urban transition to occur before the vital transition, and urban growth was actually slightly hindered in the first half of the 20th century because of a ‘demographic sink’ due to fertility after the vital transition (the lowest urban-rural natural growth difference, -0.32%, was reached in 1926).

It is difficult to say if the cases of Sweden and Belgium are representative of their respective regions (Northern Europe, Western Europe) but their comparison show contrasted effect of the vital transition on urban transition. In both cases, migration is the engine of urban transition and its contribution precedes that of vital transition, if any. Reduction in mortality in urban areas has evidently unleashed urban transition in 19th century Sweden. Reduction in fertility in urban areas only slightly reduced the speed of urban transition in 20th century Belgium.

These analyses on Sweden and Belgium appear to go counter to the de Vries argument that spatially-differentiated vital transition led to urban transition, or to use Dyson’s words that “shifts in the structure of employment are perhaps better seen as resulting from the demographic [vital] processes that bring about urbanization, rather than urbanization being seen as resulting from shifts in the structure of employment as a result of economic growth” (Dyson 2011). After replacing migration at the centre of urban transition and giving natural growth a role by default rather than at the forefront, Zelinsky’s hypothesis of mobility transition cannot be so easily dismissed. As a consequence, the role that the shift in employment structure takes in the urban transition should be re-examined.

**Discussion: spatial hierarchies in demographic transition**

When analysing the relation between two (or more) sub-populations in space, taking a political economic perspective on demographic and urban transitions is needed to consider the attraction that some space (e.g. urban areas) has on population of other spaces. Indeed, the theory of demographic transition takes its origin in the political economy of the relations between
natural resources and population, between production and reproduction. The relation between contemporary demographic transition and world economic development cannot be ignored. Following the tradition of historians that examined long-term, intertwined economic and social trends, Chase-Dunn (1998, pp. xiv-xvi) identifies several structural constants of the world-system and we will use them in our interpretation of the relation between urban and demographic transitions. Capitalism is an obvious one, and we take for granted that the reader knows the general characteristics of this mode of production well enough. More relevant to the demographic transition are two other constants of the world-system: interstate competition and centre-periphery hierarchy. These constants justify the use of the hierarchical dependence concept. Economic power is concentrated in space in core countries and cities, by way of accumulation of physical and human capital. Mortality and, more importantly, fertility decline in these centres of power indicates that the quality of human resources prevails on their quantity. These centres are the forefront of demographic transition. Migration to urban areas can be considered as a way to re-allocate human resources in the centres of economic power or to control natural resources exploitation and trade (e.g. mining cities, ports...).

In our urban-rural analysis of demographic interplay, hierarchically-defined sub-populations are attributed core and peripheral roles. Whatever the population at stake, the choice of the space must be such that at least two sub-spaces must be identified in a hierarchical way. The urban-rural divide is an obvious choice in 19th century Europe, but we will show below that for later times other hierarchies might be more relevant.

---

3 The world-system theory, inspired by the Annales group of historians around F. Braudel (1958), was incepted by I. Wallerstein (Hopkins and Wallerstein, 1982, Wallerstein, 2004) and later developed by C. Chase-Dunn (1998) who analysed urban hierarchies in history as an indicator of centre-periphery relations and evolutions.
With the consideration of hierarchies comes back the fundamental question: Is economic development the cause or the consequence of demographic transition? Hierarchical dependence implies that the economic transition to capitalism in the 18th and 19th centuries in Europe should be reinstated as the underlying cause of the demographic transition operating through the hierarchical organization of the population in the national space. Drawing again on the examples of Sweden and Belgium, we agree with Dyson that urbanization is the result of interactions between demographic movements (births, deaths, migrations). However, the central role of migration in the urbanization process shows how sensitive the whole vital and urban transitions are to economic relations between urban and rural areas. Certainly the reduction of urban mortality had a triggering effect on both urban and demographic transitions in Sweden, but questions remain: why did it occur in urban areas, and what caused this reduction in mortality in the first place?

The fact that the reduction of mortality occurred in urban areas which labour markets were also attractive cannot really be explained without reference to the economy. This is not to say that the cause of this change in demographic behaviour can readily be found in economic variables (say, industrial production, GDP, investment...), but rather that a change in the mode of production, with the necessary technological innovations to increase productivity, is a necessary and underlying condition for demographic transition to occur. This is also to say that which demographic behaviour triggered the whole transition is not constant in history. It depends on when, where and how economic changes occurred, i.e. it depends on the type of spatial hierarchical dependence at stake. In other words, the moment the transition starts in history induces different transitional modalities. In some period and place reduction in urban mortality may have triggered demographic transition; then and elsewhere it might be rural fertility, international migration, or any other combination of demographic behaviour in time and space. To note, the infamous urban penalty has not been observed at the outset of urban development in less developed countries (Gould 1998). For instance in Africa urban under-5 death rates have been lower than rural death rates and the urban penalty only re-emerged late in the twentieth
century as a significant phenomenon in the poorest urban neighbourhoods of these countries (Bocquier et al. 2011; Garenne 2010; Van De Poel et al. 2009).

Another justification for considering economic development as the underlying cause of demographic transition is found in the analysis of variations and discontinuities in the demographic and urban transitions across the world. Among the constants of the world-system identified by Chase-Dunn (1998)\(^4\), one is particularly relevant to the analysis of demographic transition: the existence of economic and political oscillation and cycles. Economic and political transitions not only vary in space and time but they are also not linear. Wars, political events, epidemics, natural disasters are as many reasons that prevent, disturb and sometimes accelerate the expected course of these transitions. These major disruptions, often observed at the global scale, may be analysed as either the result of hegemonic competition (world wars, civil wars, coup d’état, revolutions, etc.) or the failure of the society to cater for the unexpected or the unplanned (HIV/TB epidemics, earthquakes, etc.) in a highly competitive struggle for the global control of human and natural resources. The analysis of cycles in the economic and political context should help to identify cycles in the demographic domain, i.e. to distinguish stall from saturation, temporary from lasting changes.

For all these reasons, while we broadly agree with D. Reher framework that relates demographic transition, human capital formation and social and economic change (Reher 2011), we question the reduction of the economic system to a mere contextual role. If one is ready to admit hierarchical dependence and the considerable variations in vital and urban transitions, then the mode of production appears to have a long-run, necessary and underlying, although not

\(^4\) The author identifies latent trends that are remotely relevant to demographic transition: expansion of the market economy, state formation, monopoly growth, transnationalization of capital, dominance of capital over labour, proletarianization, rise of international inequalities, international political regulation (Chase-Dunn 1998, pp. xv-xvi).
deterministic, influence on demographic transition. Without referring to change in the mode of production, neither the advent of the demographic transition in developed countries nor the speed and form that it took in less developed countries can be explained. Actually, the hierarchy in economic development is closely associated with long-term inequalities in mortality, fertility, migration trends and the resulting urbanization across societies or countries.

This is not to exclude feedback effect of demographic transition on economic development such as the ‘demographic dividend’ produced by a favourable age dependency ratio (Bloom et al. 2003). This is not to say either that economic development may occur without demographic transition. On the contrary, the macro relations between development and population dynamics are better viewed as part of the expansion of the capitalist mode of production. As a major technological change, it played a role comparable in scale to that of Palaeolithic and Neolithic revolutions in human history. To note, these first two technological revolutions also brought about demographic transitions, whereby the human population increased roughly from 600,000 to 6 million circa 40-30,000 BC and from 6 million in 10,000 BC to 110 million at the beginning of the Christian era (Biraben 2006; Bocquet-Appel 2008; Bocquet-Appel and Bar-Yosef 2008).

Under hierarchical dependence principle that reinstates the importance of the change in the mode of production, the role of the urban transition becomes clearer. Urban and rural areas are hierarchically organized in an increasingly complex way as part of the transformation in the mode of production. Urban areas obviously existed before this transformation but it is only when they serve a new purpose that they developed beyond the limits in which they were constrained before. As in the Swedish example, urban development was limited by a number of economic constraints, chiefly the supply of production and reproduction inputs like wood, water, food… (van der Woude et al. 1990), as well as by demographic variables (excess urban mortality). Urban transition is the spatial manifestation of the changing relation between the demographic and economic systems. It is not so much that “urbanization […] is an inevitable outcome of the
demographic [vital] transition” (Dyson 2011, p.51, my emphasis), it is rather that urban and vital transitions are two processes which underlying cause of change originates in the transformation of the mode of production.

Hierarchical dependence has also a consequence on scaling the analysis. Countries may not be always the relevant spatial unit to analyse demographic transition occurring in the less developed countries past and present. As Zelinsky (1971) reminds us, transition modalities depend on when and where the transition started, and transition starts from a core and extends to the semi-periphery or periphery. Therefore, it is important to adapt the spatial dimension to each particular analysis. For example, the urban-rural divide echoes well the economic hierarchy at the beginning of the demographic transition as experienced in Europe. However, if the focus of the analysis is on more recent demographic transitions, the populations $i$ and $j$ can be defined along international dimension as sets of core-periphery spaces, e.g. Mexico-USA, Europe-Northern Africa, etc. By extension, spatial demographic equations can be generalized to $\times$ equations for a more complex hierarchy of three or more spatial and hierarchical entities.

Urbanization is a major phenomenon in the demographic transition observed at country level but to understand the full extent of the phenomenon other hierarchical organizations of space may be relevant at the international level, such as sub-continental, continental or multi-continental spaces. In other words, to analyze such demographic and urban transitions as global phenomena, it is necessary to depart from a methodological nationalism that restricts analysis to state boundaries. Systematic comparison of national transitions would not even suffice. A change of scale, from the national to the global, is necessary when it comes to analyze the whole history of demographic and urban transitions. This is how the apparent paradox of the so-called ‘urbanization without economic growth’ can be explained in sub-Saharan Africa or elsewhere. The economic growth that urban areas of developing countries help to generate is to be found elsewhere, i.e. in developed countries. Urban areas will grow in peripheral countries inasmuch as
they can serve core countries’ economic development. The transition in peripheral states started when the transition was well under way in center states and even in semi-peripheral states.

Conclusion

This paper is an attempt to reconsider the relationship between vital and urban transitions. Re-examining empirical evidences on Sweden and Belgium, we showed how a set of hierarchical demographic equations help in the macro-analysis of demographic transition. Results reinstate migration as the major component of urban transition. Migration is a necessary condition for urban transition, but is not always sufficient: Sweden is an example where a faster decline of mortality in urban areas as compared to rural areas made urban transition possible; Belgium is an example of negligible effect of vital transition on urban transition. Analysis of more historical series from other developed countries will be needed to generalise these results on 19th century Europe.

The explicit introduction of space and migration justifies adopting a political economic perspective on demographic and urban transitions. Changes in the economic sphere are reinstated as the underlying or exogenous cause of demographic changes acting through the hierarchical organization of populations in space. Analyzing economic and political hierarchies at international level may also help explain developing country’s demographic and urban transitions in the 20th century.

References


Figure 1: Urban-rural difference in mortality, fertility, migration and total growth in Sweden (1750-1960)

Source: our own computation using original crude birth and death estimates by urban and rural areas compiled by Dyson (2011).

Note: Estimates assuming constant 9.8% urban population before 1820 are marked with large dots.

Legend: URGD=urban-rural growth difference, URNGD=urban-rural natural growth difference, URMGD=urban-rural internal net-migration growth difference, URIGD=urban-rural international net-migration growth difference, URCDRD=urban-rural difference in crude death rates, URCBRD=urban-rural difference in crude birth rates.
Figure 2: Urban-rural difference in mortality, fertility, migration and total growth in Belgium (1840-1964)

Source: our own computation using original crude birth and death estimates by urban and rural areas.

Legend: URGD=urban-rural growth difference, URNGD=urban-rural natural growth difference, URMGD=urban-rural internal net-migration growth difference, URIGD=urban-rural international net-migration growth difference, URCDRD=urban-rural difference in crude death rates, URCBRD=urban-rural difference in crude birth rates.