Proceedings of Statistics Canada Symposium 2018
Combine to Conquer: Innovations in the Use of Multiple Sources of Data

Projecting the level of literacy using a microsimulation model

Samuel Vézina[1]

Abstract

This article provides an overview of the module used to project the literacy level of adults (aged 25 to 64 years) using the microsimulation model developed at the *LSD-C* (Laboratoire de Simulation Démographiques-Canada). It presents the method used to project the literacy level of the Canadian population and highlights the analysis of determinants of the population's literacy level, the results of which serve as literacy projection parameters. It discusses some results of projections and makes comparisons with Austria, where this method of literacy projection has also been used. Finally, it explains briefly how the literacy level projections open up a new avenue for analyzing the phenomenon of overqualification of the workforce.

Key words: Literacy; Projections; Microsimulation; Immigration; Canada.

1. Introduction

1.1 Learning to read versus reading to learn

The concept of literacy, as defined in this research, is broader than the notion of being able to read and write, although both indicators are rendered in English as "literacy." Literacy (the ability to read and write) refers to language learning and gives rise to the notions of "absolute illiteracy" and "functional illiteracy" (Wagner 2002). Literacy skills refer to the "ability to understand, evaluate, use and engage with written texts (printed and digital) in order to achieve one's goals, to develop one's knowledge and potential, and to participate in society" (OECD 2013; Statistics Canada 2013). These skills are highly valued on the labour market, as they are required for any skilled position and because they indicate that more complex skills have been mastered (Carey 2014; Levy 2010).

With the advent of knowledge societies and knowledge-based economies in the most developed countries, the issue is no longer one of learning the language, but rather one of using the language (from learning to read to literacy). Indeed, there is interest in the increasingly important role of knowledge and literacy skills in productivity, innovation and social cohesion. In this regard, the Organisation for Economic Co-operation and Development (OECD) has developed a measure of literacy that makes it possible not only to identify illiterate persons, but also to gradually rank individuals based on their effectiveness in using information to function within society and the economy. This measure is the Programme for the International Assessment of Adult Competencies (PIAAC),[2] a survey that measures individuals' literacy skills by conceptualizing them along a continuum from 0 to 500 points (de Vries 2009).

Initial analyses of the PIAAC data reveal that a number of factors, such as age and level of education affect literacy level (OECD 2013, 2016). Moreover, there is a significant gap between immigrants and non-immigrants who have an equivalent level of education, all else being equal (Xenogiani 2017). Studies in Canada have shown that the literacy level of immigrants is generally lower than that of native-born Canadians (Statistics Canada 2013), although their level of educational attainment is higher (Statistics Canada 2017b). It is essential to understand the links between

1. Samuel Vézina, Institut national de la recherche scientifique, 385 rue Sherbrooke Est, Montréal, Quebec, Canada, H2X 1E3 (samuel.vezina@ucs.inrs.ca). The results of the analyses presented in this paper were generated as part of the author's doctoral research, whose thesis was supervised by Alain Bélanger, professor of demography at the Centre Urbanisation Culture Société de l'INRS (Institut national de la recherche scientifique).
2. For more information, consult "Assessment of Adult Competencies (PIAAC)" on the website of the Organisation for Economic Co-operation and Development (OECD): http://www.oecd.org/skills/piaac/ (OECD 2014).

literacy level, immigration status and economic and social well-being, given the intensification and the ethnocultural diversification of the increase in migration to Canada (Statistics Canada 2014, 2017a).

1.2 Projecting the level of literacy using microsimulation

This research contributes to the advancement of knowledge on the impact of sociodemographic changes on the literacy level of the Canadian population. It summarizes the method used to project the literacy level of Canadians aged 25 to 64 years using a microsimulation model, developed specifically to account for Canada's ethnocultural diversification.

Microsimulation projection models are indispensable tools used to account for the new demographic regime, and for its consequences with regard to the population's ethnocultural diversification (Bélanger, Sabourin, Marois, Van Hook, and Vézina 2019). Using both the age and the gender of the simulated individuals, the microsimulation models can simultaneously project, for example, the level of education, the language, the ethnocultural characteristics of the individuals, while taking into account the fertility and mortality differentials that can be attributed to these diverse characteristics. Moreover, these models are dynamic, as the risk faced by the individuals is modified as the characteristics evolve (van Imhoff and Post 1997).

This research uses the *LSD-C* (Laboratoire de Simulations Démographiques-Canada) microsimulation model. This model projects the Canadian population by demographic variables (age, gender, place of residence, place of birth, generation status and immigration status), ethnocultural variables (mother tongue, language spoken most often at home, knowledge of official languages, visible minority group, religion) and socioeconomic variables (education, labour force status). The model uses a 2011 base population. During the simulation, the model allows for the evolution of the characteristics of the simulated cases. It also allows for intergenerational transfer of certain attributes from mother to children, a feature not found in traditional methods of component projections.[3] The reference scenario for the *LSD-C* model combines the most plausible assumptions about the future evolution of the Canadian population. The reference scenario uses the same demographic assumptions as a medium-growth scenario (M1) of Statistics Canada's most recent projections (2015).

2. How the literacy level projection module works

LSD-C therefore allows the generation of population projections for Canada using demographic, ethnocultural and socioeconomic variables. A projection module is grafted onto this model and a literacy score is imputed based on the individuals' characteristics. The score evolves throughout the simulation and is re-calculated each time the individuals' characteristics change. If, for example, an individual graduate from university, the literacy score is immediately re-calculated to account for this new characteristic. The score is also re-calculated on each birthday, when the individual moves to another region, etc.

The module calculates a literacy score for every individual aged 25 to 64 years; the score is assigned using separate parameters for the Canadian-born population and for immigrants. These parameters are kept constant throughout the projection. In other words, the relations between individuals' literacy scores and their characteristics (for example, the net age effect, or the effect of level of education) are reproduced unchanged in the model throughout the simulation. Consequently, the simulations keep the difference between the literacy level of native-born Canadians and that of immigrants constant, as was observed in the 2012 PIAAC data; the same applies to the difference between levels of education.

While it is possible, and even desirable, to close these gaps, the available data suggest that there is a certain consistency in the effects of the variables on literacy level over time. Regression analyses of the three most recent quantitative surveys on the literacy level of adults in Canada[4] have shown that the factors that determine literacy level have varied little over time (Vézina 2018a). The projection results generated by *LSD-C* show rather how future sociodemographic

3. For a detailed description of the content, methods and assumptions of the, *LSD-C model*, see Bélanger et al. (2018).
4. We refer to the 1994 ALSS, as well as the 2003 IALSS through to the 2012 PIAAC surveys.

changes, especially the increase in ethnocultural diversity and in the increase in the average level of education, can affect the literacy level of the population.

3. Parameters of the literacy level projection module

The parameters are derived from analyses of the 2012 PIAAC survey data. Given the specifications of the sampling plan for this survey, there are no parameters to impute a literacy score for non-permanent residents; residents of the northern territories are also excluded, as are individuals 65 years and over and school-age persons.

The parameters included in *LSD-C* correspond to the linear regression coefficients of the natural logarithm of the literacy score. The regressions are stratified by immigrant status (Canadian-born versus immigrants). The regression models contain several independent variables: gender, age, province of residence, level of education, knowledge and use of Canada's official languages and labour force status.[5] Four other variables are used to describe characteristics unique to immigrants: age at immigration, number of years in Canada, country of birth and country in which the individual attained the highest level of education.

The regression analyses show that level of education and, to a lesser extent, mastery of the official languages are the most powerful determinants of the literacy level of Canadian adults. Stratifying analyses by immigration status makes it possible to highlight the importance of variables specific to immigrants' literacy level. For example, the analyses show that the age at immigration significantly affects the literacy level of immigrants (Vézina, Bélanger, Sabourin and Marois 2019). The microsimulation models take this dimension into account during the imputation of the literacy score.

The factors that determine literacy level are numerous and varied. Our analysis of determinants is limited to the dimensions measured in the surveys of adult skills in Canada and which are projected in the microsimulation model. Consequently, some pertinent variables such as literacy practices in the home or at work, immigration status (refugees, economic immigrants, family class) are omitted from the analyses.

We also explored in detail the idea of including a parameter to account for the cohort effect in our literacy projection module. Indeed, recent literature reveals that the literacy level of adults in Canada is declining from one cohort to the next, *ceteris paribus*, suggesting that there is a negative cohort effect in Canada (as well as in other countries such as the United States and Norway) (Barrett and Riddell 2016; Green and Riddell 2013; Murray, Binkley, and Shillington 2016; Paccagnella 2016; Willms and Murray 2007). In our analyses, we were able to reproduce the results published in the literature using the same methodology (synthetic cohorts) and the same data from adult skills surveys in Canada. Secondly, our in-depth analyses revealed that while age has a negative effect on the literacy score, there is also a significant "period" effect that can only be explained by a change in the instrument used to measure literacy level. In other words, our analyses reveal that there is a significant effect due to the changes to the measuring instrument. If omitted from the analyses, this effect is captured by an apparent cohort effect, but is potentially misleading (Vézina 2018a). As a result, the projection module does not contain parameters to account for the cohort effect which is, in fact, more ambiguous than the literature indicates. By inserting a parameter to account for the age effect on individuals' literacy level, the effect of time on skills is quickly taken into consideration in the projection module.[6]

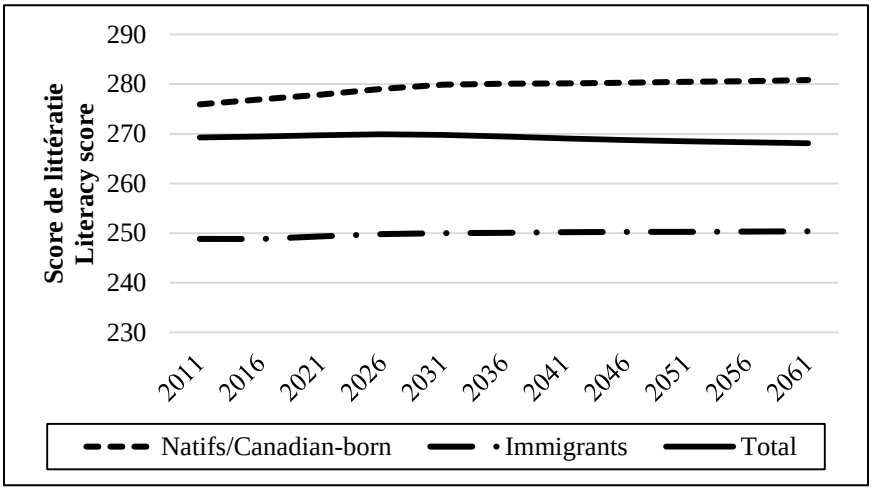
4. Some projections results

Using the projections generated by *LSD-C* and its literacy level projection module made it possible to reproduce the picture portrayed by the 2012 PIAAC data on the literacy level of the adult population in Canada. The reference

5. For more information on the detailed categorization of the diverse independent variables, see Vézina (2018b).
6. Moreover, in a similar analysis of 13 OECD countries, including Canada, it was asserted that: "*The estimated age effects from the pooled regressions are quite similar* […], *implying that the simple approach of only looking at the cross-sectional age-proficiency profile and interpreting it as "age effect" would probably not lead to significant biases*" (Paccagnella, 2016, p. 22).***

scenario revealed that there was a slight decrease in the average literacy level of the population between 2011 and 2061, which can be attributed to future sociodemographic changes. Figure 4-1 also shows the gap between immigrants and native-born Canadians with regard to average literacy level.

Figure 4-1
Projection of the average literacy level for the entire population aged 25 to 64 years and by immigration status (Canadian-born versus immigrant), Reference Scenario of LSD-C model, 2011-2061, Canada



The in-depth analyses of these projections reveal that the projected increase in the immigrant population (which has a lower average literacy level than the Canadian-born population) counter-balances the positive effect of the projected increase in the level of education of the Canadian-born population with regard to the average literacy level.

The LSD-C model and its literacy projection module made it possible to study two aspects of the impact of immigration on the literacy level of the Canadian population: the impact of the number of immigrants (volume), and the impact of the characteristics of the immigrants (composition of the immigrant population). In order to isolate the volume of immigration on the projection results, two hypothetical scenarios were created. The first **(0_IMM)** simulates a complete halt to immigration (zero immigration for the entire projection period. The second scenario doubles the volume of immigration **(2_IMM)** compared with the reference scenario (**REF**).

Figure 4-2 shows that in the zero immigration scenario (**0_IMM**), the literacy level increased by 10 points to 280 between 2011 and 2061, due mainly to the increase in the level of education of the population. When the immigration rate is doubled (**2_IMM)**), the average literacy score of Canadians aged 25 to 64 years in 2061 is 262.

Figure 4-2
Projection of the average literacy level for the total population aged 25 to 64 years and by assumptions related to the volume of immigration, *LSD-C* Model, 2011-2061, Canada

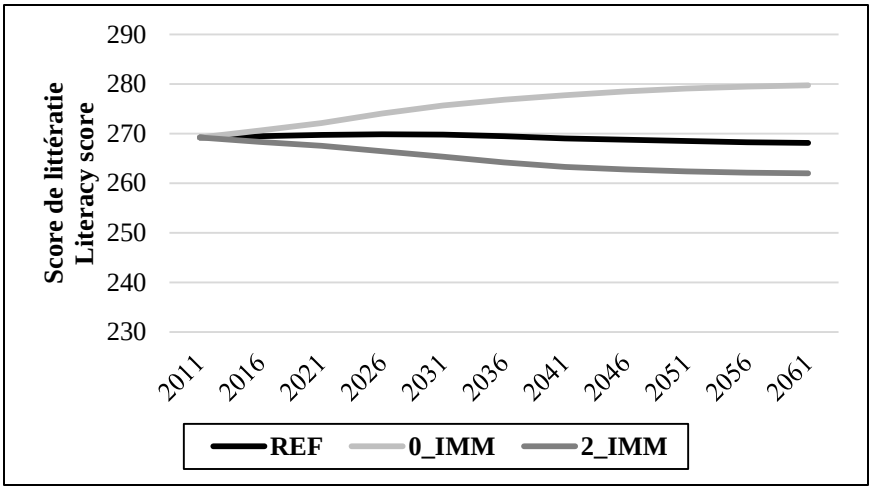


Figure 4-2 shows only the projected average literacy level of the Canadian population. The size of the total population (aged 25 to 64 years) projected in 2061 varies considerably from one scenario to the next. In the zero immigration scenario, the population in 2061 is 12 million, while it is 39.4 million in the **2_IMM**, scenario, registering a decrease of 36.4% and an increase of 108.9% respectively when compared with the initial level in 2011. By contrast, the projected average literacy level varies much less. In fact, the average literacy level increases by 3.9% in the zero immigration scenario (**0_IMM**) and decreases by 2.7%in the double immigration scenario (**2_IMM**). In other words, the projected population is more likely to be affected by an increase in immigration than the population's average literacy level.

As a result of additional projections carried out by Vézina, Bélanger, Sabourin, and Marois (2019), the simulations show that by implementing a policy for the selection of immigrants that would modify the distribution of new arrivals received annually, it would be possible to avoid the slight projected decline in the literacy level. More precisely, the projections show that by welcoming younger, better educated immigrants who have a greater mastery of Canada's official languages, there is likely to be a slight increase in the literacy level. In order to achieve the same projection results without changing the composition of the group of new arrivals, there needs to be a reduction in the total number of immigrants accepted in Canada between 2011 and 2061 during the simulation process. This reduction would amount to 3.2 million individuals or approximately 20% of the volume of immigration (Vézina, Bélanger, Sabourin, and Marois 2019). It is important to note that the fact that the OECD country with one of the smallest gaps between immigrants and the non-immigrant population is Australia, which has "extremely strict selection criteria with regard to immigration, requiring excellent mastery of English" (Xenogiani 2017, p. 3).

5. International Comparison

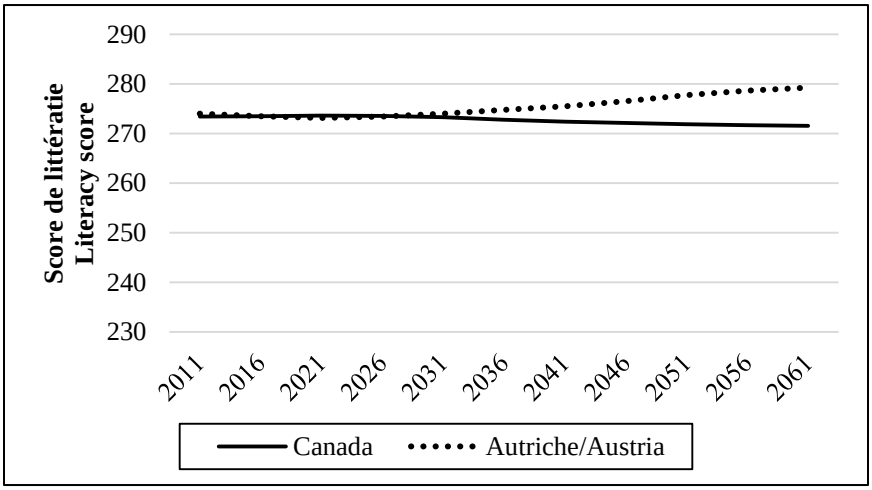
A microsimulation model similar to *LSD-C* has been developed for Austria. This is the *POB* (*Projektionsmodell der Österreichischen Bevölkerung*) model, onto which a literacy projection module has been grafted using the method described in this paper.[7] When the current government targets on immigration are used to set parameters for the simulations, the models for Canada project an increase of 27% in the size of the labour force, estimated to reach 20 million individuals by 2061. In Austria, the labour force is expected to be 4 million in 2021, and to decrease to 3.65 million by 2061 (Vézina and Bélanger 2019a).

The projections generated by *LSD-C* and *POB* show how Canadian policy, focusing on a relatively high immigration rate, supports the sustainable growth of the labour force over a 50-year period, at the expense of a constant average

7. For a detailed description of the content, methods and assumptions of *POB* model, consult Vézina (2018a). As in the case of the *LSD-C* model, the *POB* model combines the most plausible assumptions about the evolution of the Austrian population. The reference scenario uses the same demographic assumptions as the reference scenario of Statistik Austria's most recent projections (2016).

literacy level, or even a small decline in said level. The average literacy level of the labour force aged 25 to 64 moves from 273.4 to 271.5 between 2011 and 2061. By contrast, the size of the Austrian labour force in 2061 will be comparable with that of 2011 (zero growth over 50 years), while the literacy level will have increased from 274 in 2011 to 279.3 in 2061 as a result of decreased ethnocultural diversification and a more significant increase in the level of education of the population (figure 5-1). The two cases present different opportunities and obstacles. Vézina and Bélanger (2019a) maintain that it is important for countries to learn from each other's national policies and strategies to minimize the risks of adopting an approach that is too narrow and one-sided.

Figure 5-1
Projection of the literacy level of the total labour force aged 25 to 64, reference scenario of *LSD-C* Model (Canada) and the *POB* Model (Austria), 2011-2061



6. Some interesting applications

The projection results generated using the literacy level projection module grafted onto the *LSD-C* are different from the projections generated using the traditional component method in which one simply applies the literacy level by immigration status, gender and age.[8] The microsimulation projections constitute an advancement since they allow for an evolution in the probability of living through the demographic events throughout the simulation based on the changes in the characteristics of the simulated individuals. Furthermore, the model allows for the development of different evolution scenarios for the population as well as for the literacy level, yielding richer results for analysis. The microsimulation projection model is an important tool for analyzing the evolution of literacy in Canada that results from predictable changes in the characteristics of the population, of its consequences and of the policies that affect its evolution. The method developed for Canada can reliably be used for other countries and regions, as was demonstrated by the *POB* projection model used for Austria.

Using the literacy projection module, the operation and parameters of which have been presented in this paper, it becomes possible to conduct more finely tuned analyses of the skills and the quality of Canada's labour force. Indeed, further analyses of the 2012 PIAAC survey data have shown that individuals' level of education is not a perfect indicator of the actual skills that are valued on the Canadian job market; the gap with immigrants is even wider. Using the *LSD-C* model, which takes into account the fact that immigrants constitute a significant and growing part of the Canadian labour force, projections show that in accounting for literacy, there must be a downward revision of the projected figures for the labour force able to high-skilled jobs. More precisely, neglecting to factor the literacy level into the analyses results, for example, in an artificially inflated labour force distribution of almost 400,000 individuals occupying high-skilled jobs or management positions for 2024 (Vézina and Bélanger 2019b).

8. See, for example, projections of the Canada Council on Learning (2008).

This more in-depth analysis of the projected labour supply—in the sense that it goes beyond simple level of education in order to assess its quality—leads to results that are more in tune with Canada's projected labour force demands. The Canadian government's most recent official projections show an increase in the number of available jobs (demand for labour) at all levels, including less-skilled jobs. The results generated for the literacy level projection module presented in this paper offer some interesting applications, including shedding new light on the analysis of the phenomenon of overqualification of the labour force in Canada.

References

Barrett, G., and W. C. Riddell (2016), "Ageing and Literacy Skills: Evidence from IALS, ALL and PIAAC", OECD Education Working Papers, N° 145, Paris: OECD Publishing.

Bélanger, A., P. Sabourin, G. Marois, J. Van Hook and S. Vézina (2019), "A framework for the prospective analysis of super-diversity", *Demographic Research*, (acceptée).

Bélanger, A., P. Sabourin, S. Vézina, G. Marois, K. D'Ovidio, D. Pelletier and O. Lafontaine (2018), "The Canadian microsimulation model (LSD-C): Content, modules, and some preliminary results", Working paper N° 2018-01, Montréal: Institut national de la recherche scientifique.

Carey, D. (2014), "Combler les pénuries de compétences au Canada", OECD Economics Department Working Papers, N° 1143, Paris: Éditions OCDE.

Conseil canadien sur l'apprentissage (2008), "Lire l'avenir pour répondre aux besoins futurs du Canada en matière de littératie", Ottawa: Conseil canadien sur l'apprentissage.

de Vries, E. (2009), "Workplace Essential Skills in Policy and Practice: A Canadian Perspective", R. Maclean et D. Wilson (éd.) *International Handbook of Education for the Changing World of Work*, Netherlands: Springer, p. 2713-2730.

Green, D., and W.C. Riddell (2013), "Ageing and literacy skills: Evidence from Canada, Norway and the United States", *Labour Economics*, 22, p. 16-29.

Levy, F. (2010), "How Technology Changes Demands for Human Skills", OECD Education Working Papers, N° 45, Paris: OECD Publishing.

Murray, S., M. Binkley and R. Shillington (2016), "Reconstructing the Evolution of the American Supply of Cognitive Skills: A Synthetic Cohort Analysis", Commissioned Paper, Washington DC: National Center for Education Statistics.

OCDE (2014), *OCDE - Évaluations des compétences*.

OECD (2013), "OECD Skills Outlook 2013: First Results from the Survey of Adult Skills", Paris: OECD Publishing.

OECD (2016), "Skills Matter: Further Results from the Survey of Adult Skills", Paris: OECD Publishing.

Paccagnella, M. (2016), "Age, Ageing and Skills: Results from the Survey of Adult Skills", OECD Education Working Papers, N° 132, Paris: OECD Publishing.

Statistik Austria (2016), "Demographisches Jarhbuch 2015", Vienna: Statistik Austria.

Statistique Canada (2013), "Les compétences au Canada : Premiers résultats du Programme pour l'évaluation internationale des compétences des adultes (PEICA)", N° 89-555-X au catalogue, Ottawa: Statistique Canada.

Statistique Canada (2014), “Croissance démographique: l’accroissement migratoire l’emporte sur l’accroissement naturel”, N° 11-630-X au catalogue, Ottawa: Statistique Canada.

Statistique Canada (2017a), “Immigration et diversité ethnoculturelle : faits saillants du Recensement de 2016”, N° 98-402-X2016007 au catalogue, Ottawa: Statistique Canada.

Statistique Canada (2017b), La scolarité au Canada : faits saillants du Recensement de 2016”, Composante du produit N° 11-001-X au catalogue, Ottawa: Statistique Canada.

van Imhoff, E., et W. Post (1997), “Méthodes de micro-simulation pour des projections de population”, *Population*, 52, 4, pp. 889-932.

Vézina, S. (2018a), “Annexe: Analyse transversale et pseudo-longitudinale des déterminants de la littératie”, Dans S. Vézina (éd.) *Projection par microsimulation du niveau de littératie de la population adulte canadienne: Regard sur la population immigrante et son intégration sur le marché de l'emploi canadien*, Thèse de doctorat, Montréal: Institut national de la recherche scientifique, pp. 146-188.

Vézina, S. (2018b), “Projection par microsimulation du niveau de littératie de la population adulte canadienne: Regard sur la population immigrante et son intégration sur le marché de l'emploi canadien”, Thèse de doctorat, Montréal: Institut national de la recherche scientifique.

Vézina, S., and A. Bélanger (2019a), “Impacts of Education and Divergent Immigration Policies on the Size and Skills of the Future Workforce”, *Demographic Research*, (en évaluation).

Vézina, S., and A. Bélanger (2019b), “Literacy Skills as an Explanation for Labor Market Imbalances by Occupational Type in Canada: Microsimulation Projections for 2014-2024”, *Population Research and Policy Review*, (en évaluation).

Vézina, S., A. Bélanger, P. Sabourin and G. Marois (2019), “Literacy skills of the future Canadian working-age population: Assessing the skill gap between the foreign- and Canadian-born”, *Canadian Studies in Population*, (acceptée).

Wagner, S. (2002), “Alphabétisme et alphabétisation des francophones au Canada : Résultats de l’Enquête internationale sur l’alphabétisation des adultes (EIAA)”, N° 89-552-MIF au catalogue — n° 10, Ottawa: Statistique Canada et Développement des ressources humaines Canada.

Willms, D. J., and S. T. Murray (2007), “Acquisition et perte de compétences en littératie au cours de la vie”, N° 89-552-MIF au catalogue — n° 16, Ottawa: Statistique Canada.

Xenogiani, T. (2017), “Pourquoi les adultes immigrés sont-ils moins compétents en littératie que leurs pairs autochtones?”, Paris: Éditions OCDE.