impact canada

Summary Report

Our rose, thorn and bud:

The complexities of using tax data to measure the impact of Challenges

Government of Canada / Gouvernement du Canada

Canada

Our rose, thorn and bud: The complexities of using tax data to measure the impact of challenges.

For more information, contact:

Impact Canada, Privy Council Office
85 Sparks Street, Room 1000
Ottawa ON Canada K1A 0A3
info@pco-bcp.gc.ca

Cette publication est également disponible en français : *La complexité de l'utilisation des données fiscales pour mesurer l'incidence des défis : Notre rose, notre épine et notre bourgeon.*

ISBN: ISBN: 978-0-660-69728-4
CP22-211/2024E-PDF

Contents

Introduction

Since 2017, Impact Canada has been partnering with federal partners to leverage Challenges as an innovative financing tool to help solve complex problems for Canadians. Impact Canada Challenges differ from other traditional government funding programs as they are designed as an outcomes-based approach[1] using financial and non-financial incentives to draw innovators to direct their talents towards finding solutions to specific problems.

Since the introduction of Challenges, Impact Canada has been working on measuring their impact, as well as improving how they are designed and implemented. Guided by a clear Logic Model, Impact Canada has operationalized both quantitative research with Statistics Canada, and qualitative research, including case studies, surveys and social network analyses. To read more about Impact Canada's impact measurement approach, please see this blog.

In previous reports, we presented highlights from Stakeholder surveys and results from a social network analysis conducted in collaboration with students from the University of British Columbia, and Statistics Canada to measure the impact of our Deep Space Food Challenge.

Exploring quasi-experimental analyses

What was done?

Since 2019, Impact Canada has partnered with Statistics Canada's Centre for Special Business Projects (CSBP) to explore the use of tax data to inform on the performance of Challenge applicants and to evaluate the economic impact of Challenges using Statistics Canada's Business - Linkable File Environment (B-LFE). The B-LFE is an environment in which different administrative and survey data sources are linked to Statistics Canada's business microdata.

Quasi-experimental analyses were explored to attempt to replicate the rigour of randomized controlled trials, which is not possible to implement in a Challenge context as competition winners are not randomly assigned. To demonstrate causal effects of Challenges on outcomes of interest, data on Challenge applicants, collected by Impact Canada, were linked to the Business Register (BR). The BR is a repository of registered businesses operating in Canada maintained by Statistics Canada. This dataset was subsequently linked to tax data, from the B-LFE, to understand what types of businesses[2] participate in Challenges by assessing their business characteristics and economic performance. A non-applicant group was created by exploiting business similarities across a series of variables to conduct a comparative analysis.

This approach was two-fold. First, to understand if Challenges mobilize new talent, we profiled business applicants by geographic and business characteristics. Second, to understand if Challenges increase investments and enhance skills and capacity, we compared the business performance between challenge applicants and similar non-applicants on the following indicators:

1 Funding is rewarded on the achievement of an outcome rather than paying on expenditure or activities.

2 Only Challenge business applicants were in-scope for this analysis.

- Revenue
- Assets
- Liabilities
- Employment
- Salaries and wages
- Labour productivity
- Research and development (R&D) expenses

Two rounds of quasi-experimental analyses have been conducted. The first was conducted to assess the possibility of evaluating the success of Challenges achieving their goals. As Impact Canada launched new Challenges and the number of Challenge applicants grew since the first study, a second analysis was conducted. The second analysis aimed to expand depth and scope of the first study by including the additional years of business data and assessing the feasibility of performing a program impact analysis.

Why was this done?

Unlike surveys and other methods of gathering data, using administrative data, such as tax data, for purposes other than it's intended use enables access to high quality data to profile business economic performance. Importantly, the use of administrative data reduces respondent burden of applicants involved in Challenges from having to answer questionnaires.

The Statistic Act allows Statistics Canada to link tax data with other survey and administrative data for research purposes. This allows Statistics Canada to collaborate with government departments "in the collection, compilation and publication of statistical information, including statistics derived from the activities of those departments". The linked data offers the ability to increase the scope and depth of possible research to be conducted. For example, using tax data can open possibilities for proxy variables to measure outcomes.

Exploring a quasi-experimental analysis using the B-LFE was compelling to assess the impact of Challenges because it generates evidence on how challenge applicants perform post-Challenge from an investment perspective. It also enables the comparison of applicants to similar businesses that did not apply to Impact Canada Challenges – a feature that allows us to assess how effectively the Challenge program outcomes are being achieved.

To systematically summarize the results and learnings from this work, a rose analogy is used. The report below first summarizes what went well and the findings from the analyses (our rose), then outlines the obstacles we've faced (our thorn), and lastly discusses what we've learned and future opportunities to further develop this work (our bud).

Our rose: What we've learned so far

The first quasi-experimental analysis was conducted in 2021, which analysed 7 Challenges from 2017 to 2019. As more data became available, a second quasi-experimental analysis was conducted in 2023 on 12 Challenges from 2016 to 2021. In conjunction with the second analysis, a complementary analysis was conducted to assess the feasibility of performing an impact analysis of the Challenge program.

The results from the second analysis, as shown in **Figure 1**, illustrate that 90% of Challenge business applicants were for-profit organizations. Applicants were spread out geographically, with three out of four business applicants operating in Ontario (41%), the Prairies (18%) or British Columbia (16%).

The majority (95%) of applicants were small businesses, with less than 100 employees, and over half (59%) were young businesses who had been operating for less than 5 years, as seen in **Figure 2**.

Figure 3 illustrates that while businesses that apply to Impact Canada Challenges are diverse and operate in 20 different industries, over half (59%) of them are in the Manufacturing (22%) and Professional, scientific and technical services (37%) industries.

To inform on whether Challenge applicants were indeed able to increase investments as a result of the supports provided throughout Challenges, revenues, assets, liabilities of Challenge applicants were analysed and then compared against non-applicants.

The comparative analysis illustrated that Challenge applicants are in fact increasing investments more than non-applicants, as growth rates of revenue (104% and 50%), assets (125% and 111%) and liabilities (226% and 122%) between 2016 and 2020 were higher for applicants than non-applicants, as seen in **Figure 4**.

To inform on whether Challenges were improving skills and capacity of applicants, employment and payroll data was analysed. **Figure 5** demonstrates that Challenges are indeed enhancing skills and capacity as growth rates from 2016 to 2021 of employment (58% and 36%), wages (100% and 62%), wage per employee (23% and 15%) and labour productivity[3] (124% and 62%) were higher for Challenge applicants than non-applicants.

R&D expenditures[4] were assessed as a proxy to understand if Challenges were contributing to the development of innovative products and services. The results illustrate that between 2016 and 2020 non-applicants (28%) increased their spending on R&D expenditures more than applicants (14%), as shown in **Figure 6**.

Part of the second round of analysis included assessing the feasibility of performing a program impact analysis. The work intended to explore the types of statistical methods that could be employed for an impact analysis and the feasibility of conducting an impact analysis in the future once additional data becomes available. The feasibility study would be used to document strength and weaknesses of approaches, and suggestions on what should be done in subsequent years to facilitate a successful impact analysis.

3 Labour productivity represents the ratio of total revenue divided by the average number of employees.
4 Data coverage for R&D expenditure is often more scarcely available than other variables.

What did the feasibility analysis inform?

An assessment of the feasibility of performing an impact analysis was conducted based on the data available to determine the impact of participating in Impact Canada Challenges. That is, to answer the question: is there statistically significant evidence that participating in Impact Canada Challenges measurably improves applicants' investment and employment outcomes?

To ensure sufficient data was available pre and post Challenge, the impact analysis was conducted on 40 businesses who won at least one prize (semi-finalists, finalists, and grand-prize winners), coming from 6 Challenges between 2016 and 2021. A difference-in-differences[5] approach was used to test the effects of participating in Challenges on a set of outcome variables including, revenue, expenses, assets, liabilities, employment, wages, and labour productivity.

Results from the feasibility study showed no statistically significant evidence that participating in a Challenge had an impact on any of the dependent variables tested (revenue, expenses, assets, liabilities, employment, wages, and labour productivity). However, this work has provided the groundwork towards estimating the impact of Impact Canada Challenges in the future.

Our thorn: The limitations

The limitations when measuring the impact of Challenges

Impact Canada Challenges vary by subject-matter, intended outcomes, timelines, number of stages, prize purse, and number of applicants. This results in a complex setting for impact evaluation, especially given the variability across Challenges and applicants. This variability implies a degree of heterogeneity.

This heterogeneity among applicants is further increased as Challenges draw from diverse types of applicants such as businesses, academics, individuals, government and indigenous organization, and international applicants. Coupled with this, as the Challenge stages progress, fewer applicants advance to become semi-finalists, finalists, and grand-prize winners, which thereby reduces the sample size of applicants that are in a better position to demonstrate if Challenge outcomes are being achieved. This poses a challenge, as it makes sense to focus on applicants that have successfully progressed further as they have incrementally received more government funding than unsuccessful applicants; and have presumably had more time to commercialize and scale. To justify this notion and to generate evidence on the impact of Challenges on business performance of successful applicants, further empirically testing based on theory and past studies would need to be done.

The small sample size of semi-finalists, finalist and grand-prize winners also prevents the ability to estimate the impact on a per-Challenge basis. As a result, to assess the impact of Challenges on successful applicants, Challenges that may have different impacts at different scales are aggregated.

---

5 The difference-in-differences framework relies on assessing a difference between two differences. The first difference estimates the variation in outcome through time (pre- vs post-treatment), and the second difference estimates the difference in trends between the treated and control groups. This way, pre-existing unobservable time-invariant differences between the treated and control groups are accounted for as their effect on the outcome variable is subtracted through the first difference.

The limitations of using tax data to measure the impact of Challenges

Despite the advantages of using corporate tax data to measure the outcomes and impact of Challenges on business performance, it also poses limitations as only business applicants whose information can be linked to tax data can be used for analysis. This approach does not apply to academics, individuals, and non-business applicants, limiting program impact to only a business perspective. Hence, this further decreases the sample size and has implications on the level of precisions of statistics presented to profile business performance using tax data given data confidentiality requirements. Furthermore, some tax information is not provided by all businesses, such as R&D expenditures, which limits proxy variables available to assess if investment are made in developing innovative products and services with respect to support provided by Challenges.

The heterogeneity both at the program and applicant levels, small sample size, and inevitable data outliers – which cannot always be removed due to the already small sample size – make it difficult for a statistical model to provide evidence of impact in the context of this research and have highlighted the trade-offs between increasing precision and robustness.

These limitations coupled with varying data availability also restricts the ability to compare the impact of a Challenge against another in a given year. For the time being, point-in-time measures are obtained for Challenges both completed and in-progress, to get a sense of the impacts of Challenges.

Our bud: Lessons learned and future work

What have we learned?

Exploring quasi-experimental analyses using tax data has demonstrated to be a rigorous method of generating evidence that is adaptable to measure outcomes and impact as they unfold. Furthermore, this work has demonstrated that using administrative data to assess economic impacts of programs can also inform on larger program or policy implications.

Given that corporate tax datasets are both comprehensive and timely, this opens up opportunities for future longitudinal and cross-sectional analyses and offers the ability to increase the scope and capabilities of research if needed. However, to do this, and to overcome many of the limitations posed by the Challenges design, methodology, and current data availability, we have learned that we'll need the fullness of time. As more recent Challenges are completed and sufficient pre and post data becomes available, the number of successful applicants also increases, which can contribute to improving the robustness of statistical analyses and allows to inform on longer-term outcomes, such as economic improvements.

The partnership between Impact Canada and Statistics Canada has enabled an iterative research and analysis initiative, where we are able to build on our learnings. The collaborative work done to assess the impact of Challenges thus far has set a solid foundation and provides many opportunities to explore other quasi-experimental approaches, and to increase the robustness of analyses conducted in

upcoming years. As more data becomes available, in 2024 and beyond, we are hopeful that additional analytical opportunities will enable us to explore more research questions to further inform on outcome-based funding.

What could we explore in the future?

To improve how Impact Canada Challenge outcomes are measured, the quasi-experimental analyses could be enhanced in the future work by exploring potential linkages between available tax data, patent data, and investment open data. Additional methodologies, such as dynamic difference-in-differences models which could provide impact estimates for each year after the beginning of a Challenge, could also be explored to enhance the impact analyses.

Acknowledgments

Thank you to our partners at Statistics Canada, Xavier Bouchard-Vachon and Simon Dessureault, for the ongoing partnership and collaboration on this research. Working with a statistical agency has been helpful for Impact Canada to test and refine our approaches.

Annex

Figure 1: Breakdown of Challenge business applicants by for-profit/non-profit and by geographic region

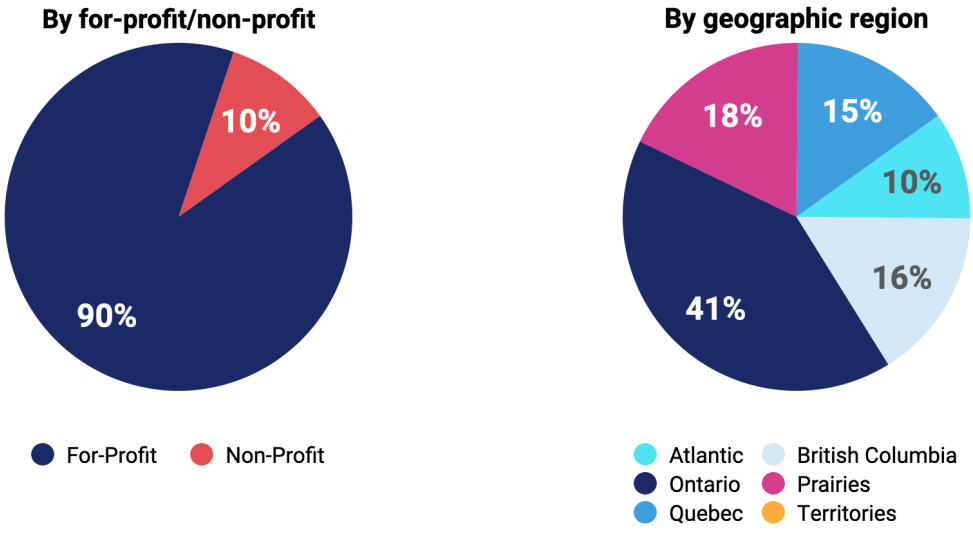


Figure 2: Breakdown of Challenge business applicants by number of employees and by number of years in operations

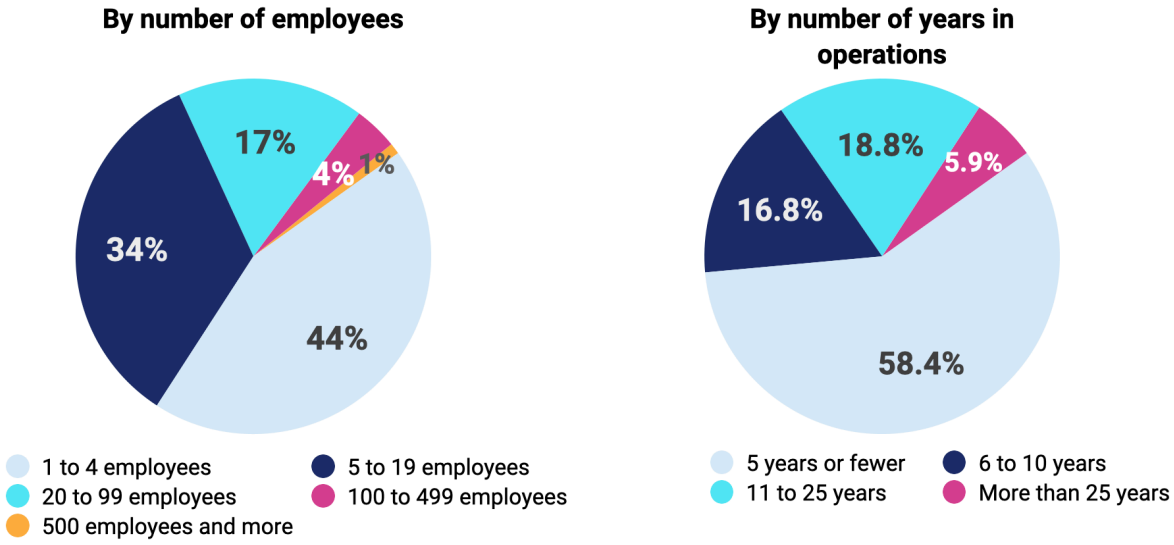


Figure 3: Breakdown of Challenge business applicants by industry of major business activity

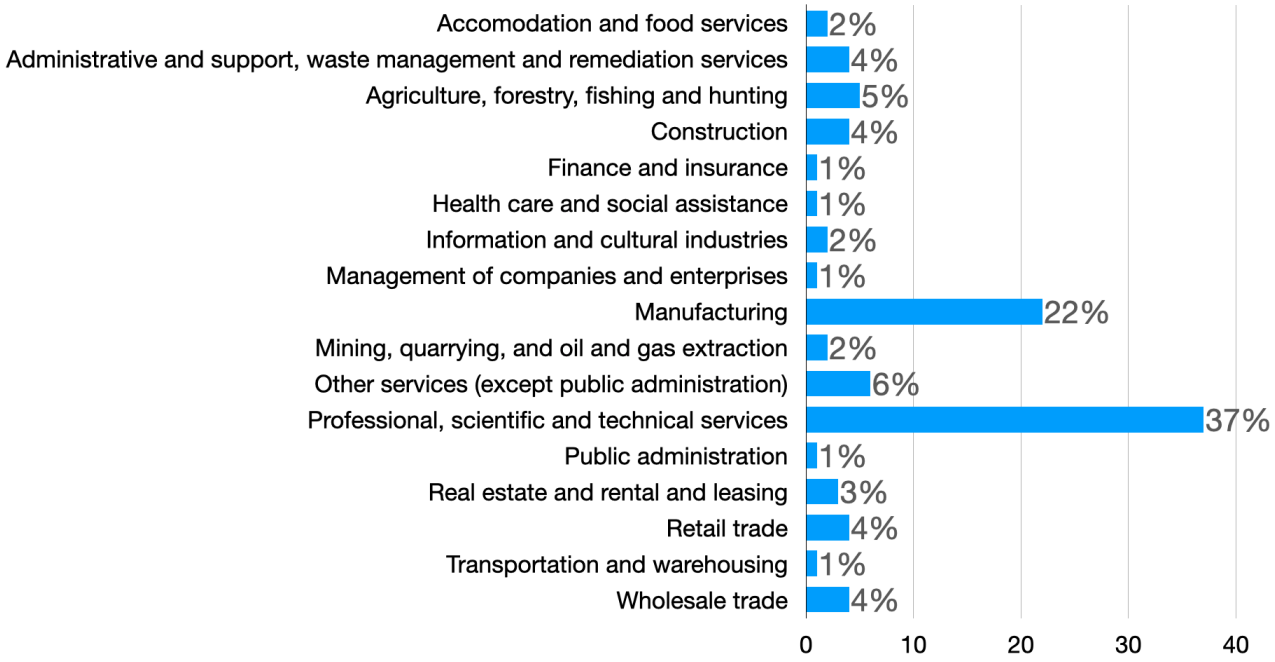
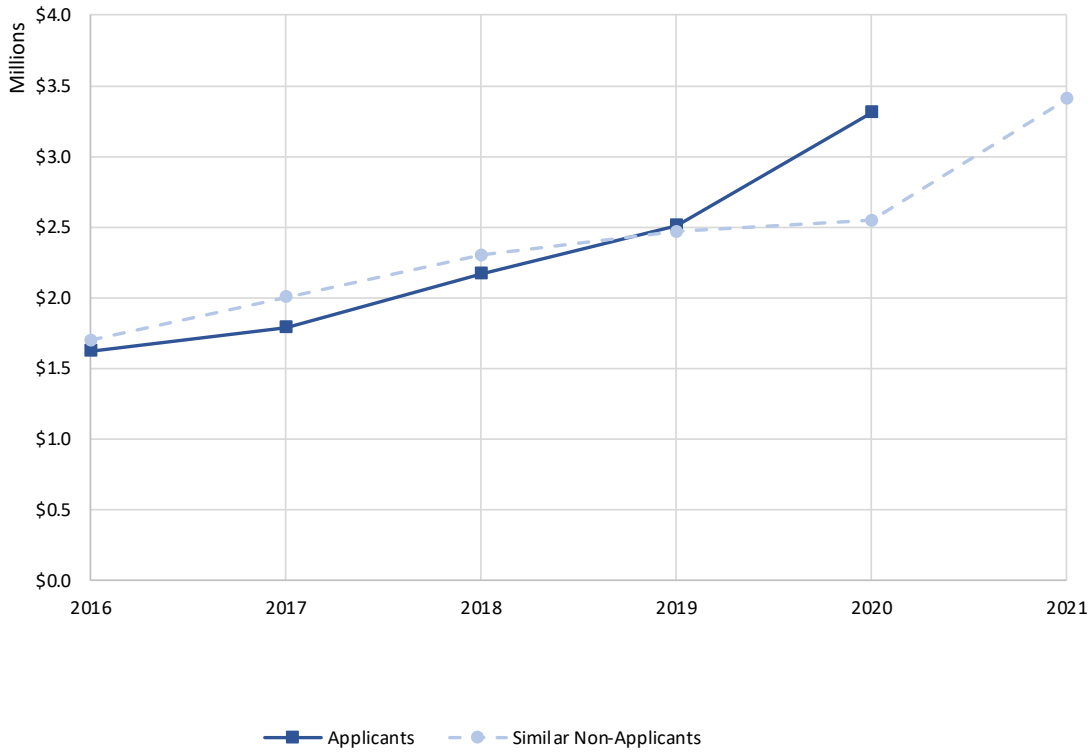


Figure 4: Revenue, assets and liabilities of Challenge applicants and non-applicants

TOTAL REVENUE



ASSETS AND LIABILITIES

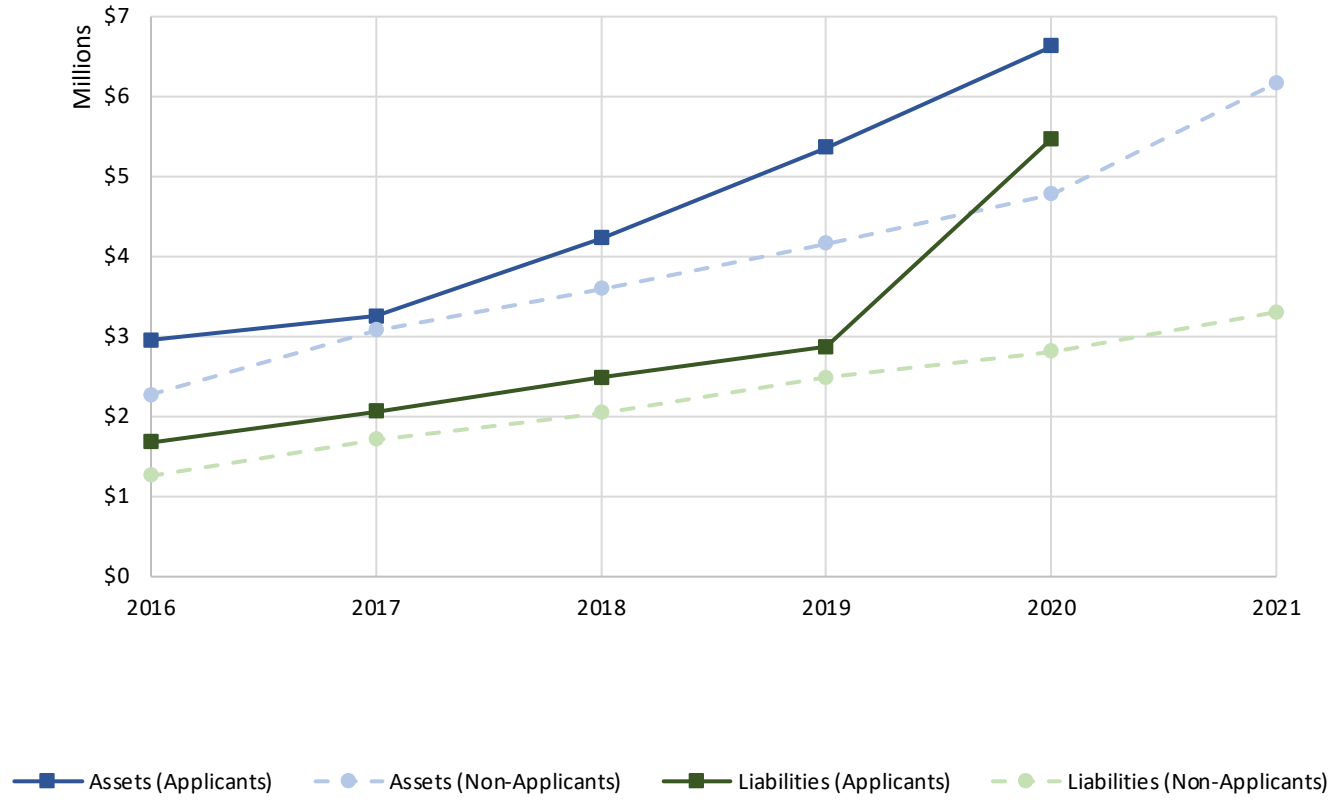
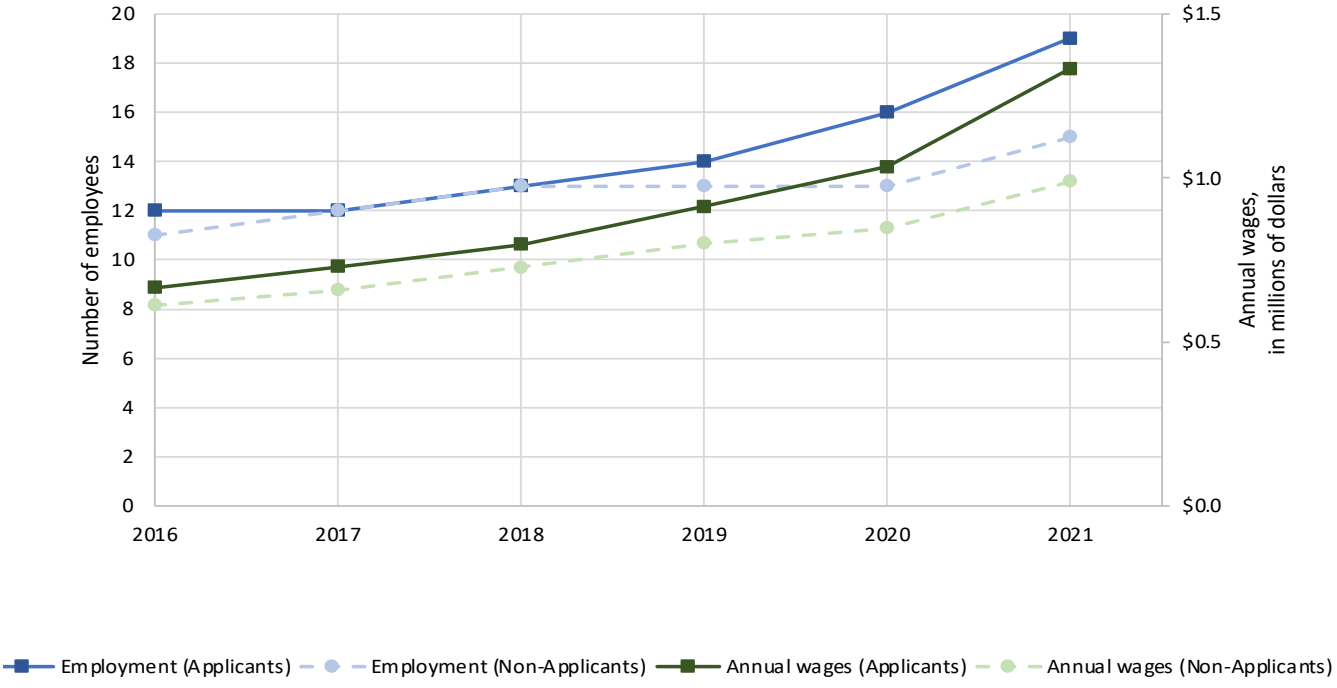


Figure 5: Employment, wages, wage per employee and labour productivity of Challenge applicants and non-applicants

EMPLOYMENT AND ANNUAL WAGES



AVERAGE WAGE PER EMPLOYEE AND LABOUR PRODUCTIVITY

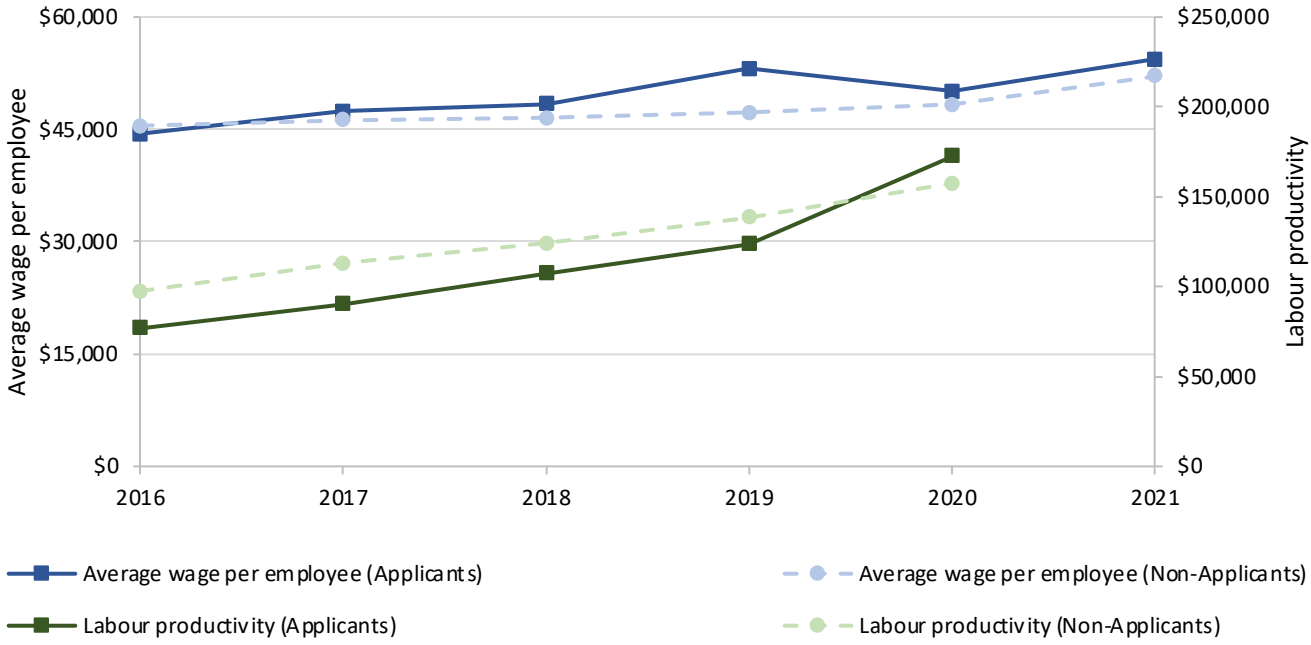


Figure 6: R&D expenditures of Challenge applicants and non-applicants

R&D EXPENDITURES