Impact assessment of AgriInnovation-Stream C and AgriInnovate: Dynamics and drivers (Phase II)

Research and Analysis Directorate
Strategic Policy Branch
Agriculture and Agri-Food Canada

The AgriInnovation-Stream C and AgriInnovate Programs

AgriInnovation-Stream C (2013-14 to 2017-18) and AgriInnovate (2018-19 to 2022-23) aimed to accelerate the commercialization phase of the innovation process in the sector using interest-free, repayable contributions for eligible innovation projects.

Table 1. Program participants, AgriInnovation-Stream C and AgriInnovate programs, 2013-2021

Program	**AgriInnovation-Stream C**					**AgriInnovate**				**Total**
Year	2013	2014	2015	2016	2017	2018	2019	2020	2021	
Number of participants	9	10	9	12	5	6	7	6	7	**71**
Average funding ($ million)	5.9	3.5	4.1	1.8	1.7	4.1	1.9	5.2	2.5	**3.4**

Note: two firms are not qualified for causal analysis, leaving 69 in-scope firms. *Source*: SDPAU, Programs Branch, AAFC.

Economic theory

Government innovation subsidies incentivize firms to invest more in such activities, thereby improving the performance of the individual firms as well as the broader economy (Afcha and Lopez, 2014; Chudnovsky et al., 2006; Masso and Vahter, 2008; Hall et al., 2009; Cin et al., 2017).

Research questions

Phase I	*Are AAFC's AgriInnovation Stream C and AgriInnovate Programs effective in improving the economic performance of the recipient firms?*
Phase II	*1. How long does it take for the program to show its impacts?*
	2. How long does the impact of the program last?
	3. What industry is the main driver of the impact?
	4. What firm size is the main driver of the impact?
	5. Which gender experiences a larger impact?

Data

Source: Business-Linkable File Environment (B-LFE) and Diversity and Skills Database (DSD).

Time frame: 2012-2022.

Information:

- Financial (income tax data): total revenues, operational expenses;
- Employment (payroll deduction accounts): number of employees, salaries and wages;
- Diversity: share of immigrant or female employees in workforce, average age of employees;
- Other: Location, research and development and gender of owner.

Table 2. Number of observations and firms in the dataset, 2005-2020

	Participants		Non-participants		Full sample	
	Firms	Obs.	Firms	Obs.	Firms	Obs.
Qualified for analysis (before matching)	69	759	103,885	723,449	103,954	724,208
Used in models (after matching)	63	516	182	1,453	245	1,969

Note: Matches could not be found for six firms, leaving 63 in-scope firms for the main models. *Source*: Statistics Canada's B-LFE.

Methodology

Objectives

- To investigate a causal link from program participation to financial performance – e.g., revenues; and
- To explore factors that affect program participation.

Approach

Step 1: Build a control group: propensity score matching
Participants are matched with non-participants of similar **observable** traits.

Step 2: Estimate the program impact
Apply two-way fixed effects (**TWFEs**) difference-in-difference (**DID**) regressions to the matched observations.

Advantages

- Matching: Assurance of random selection – i.e., no false comparisons;
- TWFEs DID: Control for some firm-specific factors – e.g., managerial talent; and
- Together, matching and TWFEs DID mostly address self-selection bias.

Note: This approach was first introduced by Heckman et al. (1997) and Heckman et al. (1998); other applications: Arnold and Javorcik (2009), Gorg et al. (2008), and Volpe and Carballo (2008).

Propensity Score Matching

1. To match participants and non-participants, we first **estimate the propensity of participation** (using a logistic regression model) as a function of firms' observable characteristics – e.g.,

   - *Exporter status* (represents whether the firms is an exporter)
   - *Years since birthdate* (represents age of business)
   - *Revenue class* (indicator of firm size)
   - *Industry category* (37 distinct categories based on their NAICS codes)
   - *Net income class* (indicator of firm size)
   - *Year* (2012-2022)

2. Participants are then matched with non-participants from the **same industry category, calendar year, and revenue class** whose **propensity score** is within a pre-specified radius caliper of that of the participant.

TWFEs Panel DID regressions

The **DID** model is estimated as a **fixed effects (FEs) panel** of the following form:

$$(1) \qquad Y_{it} = \alpha + \beta(\boldsymbol{Treatment_{it}}) + X_{it}\lambda + \gamma_i + v_{it},$$

where Y_{it} is a measure of financial performance – e.g., net income – of firm *i* in time *t*, $\boldsymbol{Treatment_{it}}$ takes the value of funding for participant firms after participation and 0 otherwise, and the set of other explanatory variables X_{it} is as follows:

- *Year* (2005-2020)
- *Industry category* (37 distinct categories based on their NAICS codes)
- *Province*
- *Year* × *Industry category* (interaction of calendar years and industry categories)
- *Year X Province* (interaction of calendar years and province)
- *Year X Province X Industry category* (interaction of calendar years, province, and industry categories)

Note: α is the constant term, β is the parameter that measures the effect of the program, λ is a set of parameters to be estimated, γ_i represents firm-specific (observable and unobservable) time-invariant characteristics or fixed effects, v_{it} reflects the remainder stochastic disturbances.

Self-selection bias: Picking winners

Problem

There may be something about the participating firms or their managers that positively affects both their likelihood of participation and their outcome – e.g., opportunistic vs. less opportunistic managers.

Implication

The estimated impact is confounded with the effect of the firms' or their managers' intrinsic abilities.

Solution

If source of bias is:

1. ***Observable characteristics*** (e.g., size, value of assets) → ***Matching***
2. ***Unobservable time-invariant*** (e.g., the manager's character) → ***FEs***
3. ***Unobservable time-variant*** (e.g., temporal changes in an industry, province, or firm's conditions)
   a) At industry level→ ***Year*** × ***Industry*** in ***regression model***;
   b) At province level→ ***Year*** × ***Province*** in ***regression models***;
   c) At industry *j* in province *k* level→ ***Year*** × ***Province*** × ***Industry*** in ***regression models***;
   d) At firm level → Two supplementary approaches – using the **not-yet-treated** and the **rejected applicants**.

Heckman et al. (1997) claim, "this estimator [matching DID] is effective in eliminating bias, especially when it is due to temporally-invariant omitted variables." Argued by Blundell and Costa Dias (2002) and stated in Görg et al. (2008): "a combination of matching and DID analysis may be a particularly suitable approach" to address self-selection bias. See **Handout 1** for more on the literature.

Program impacts

Overall impact: The program has a statistically significant impact on the financial performance of the participants.

Table 3. Estimated impact of $1 of program funding on an average participant, cumulative, over 2013-22.

	Estimated impact	Std. Err	P-value
Total revenues	**6.26**	2.73	0.02
Total expenses	**5.82**	2.69	0.03
Gross profit (loss)	**1.13**	0.31	0.00
Net income (loss) before tax	**0.45**	0.17	0.01

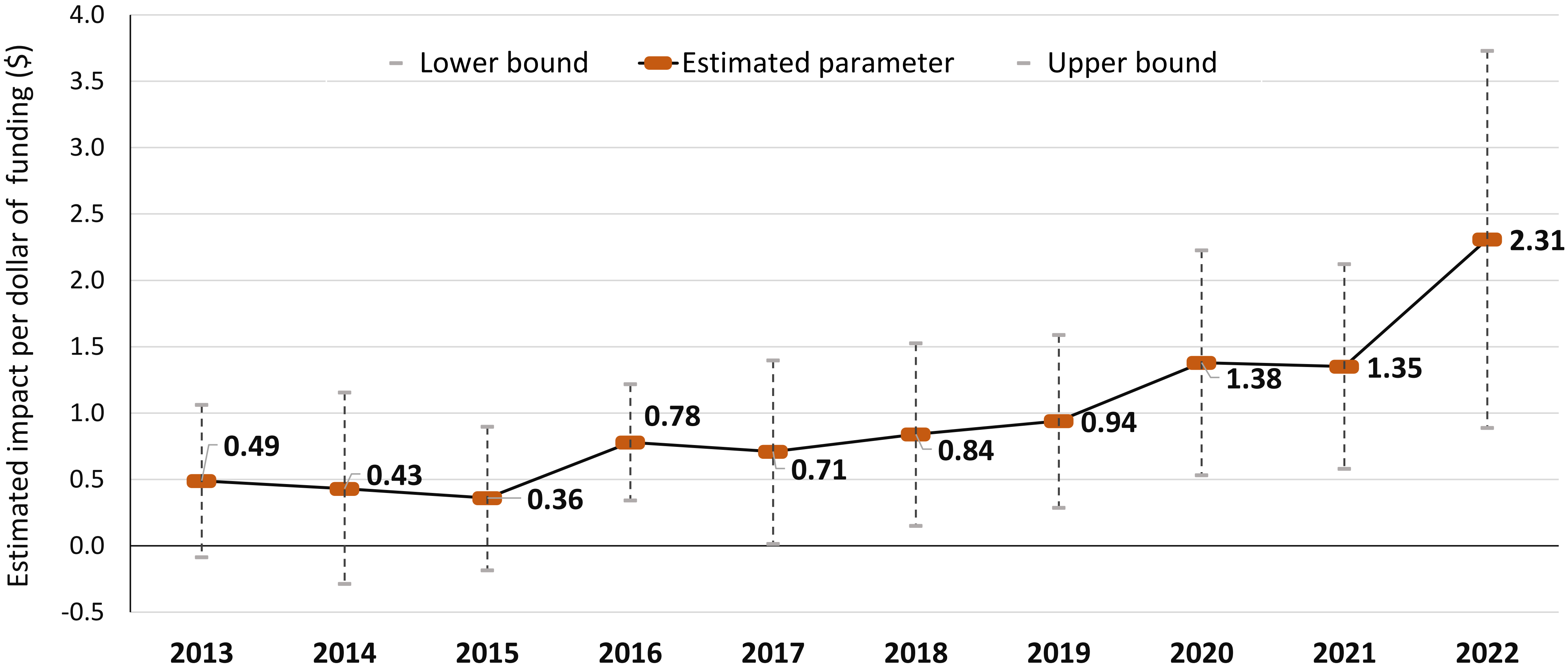
	Observations	Firms
Treated	516	**63**
Untreated	1,452	182
Total	1,968	245

Note: 1. *Total expenses*: *Cost of sales* plus *total operating expenses*; Gross profit (loss): *Total revenue* minus *Cost of sales*; and *Net income (loss) before tax*: *Total revenue* minus *total expenses*. Thus, **the difference between *net income* and *gross profit* is *total operating expenses***, which is the sum of all indirect costs such as advertisement costs, interest payments, amortization, insurance costs, etc. 2. see Slide 7 for exact model specification; *Source:* Authors' estimations.

Dynamics

Finding 1: The program impact grows over time, taking at least 3-4 years for initial impacts to appear. After this, these impacts grow over time for at least 7 years.

Figure 1. Estimated program impact on average participants' gross profit: Temporal effects and 95% confidence intervals.



Finding 2: The manufacturing sector is the main driver of the overall program impact, though the impact is largest for the wholesale sector.

Figure 2. Estimated program impact on an average participant's gross profit: Industry effects

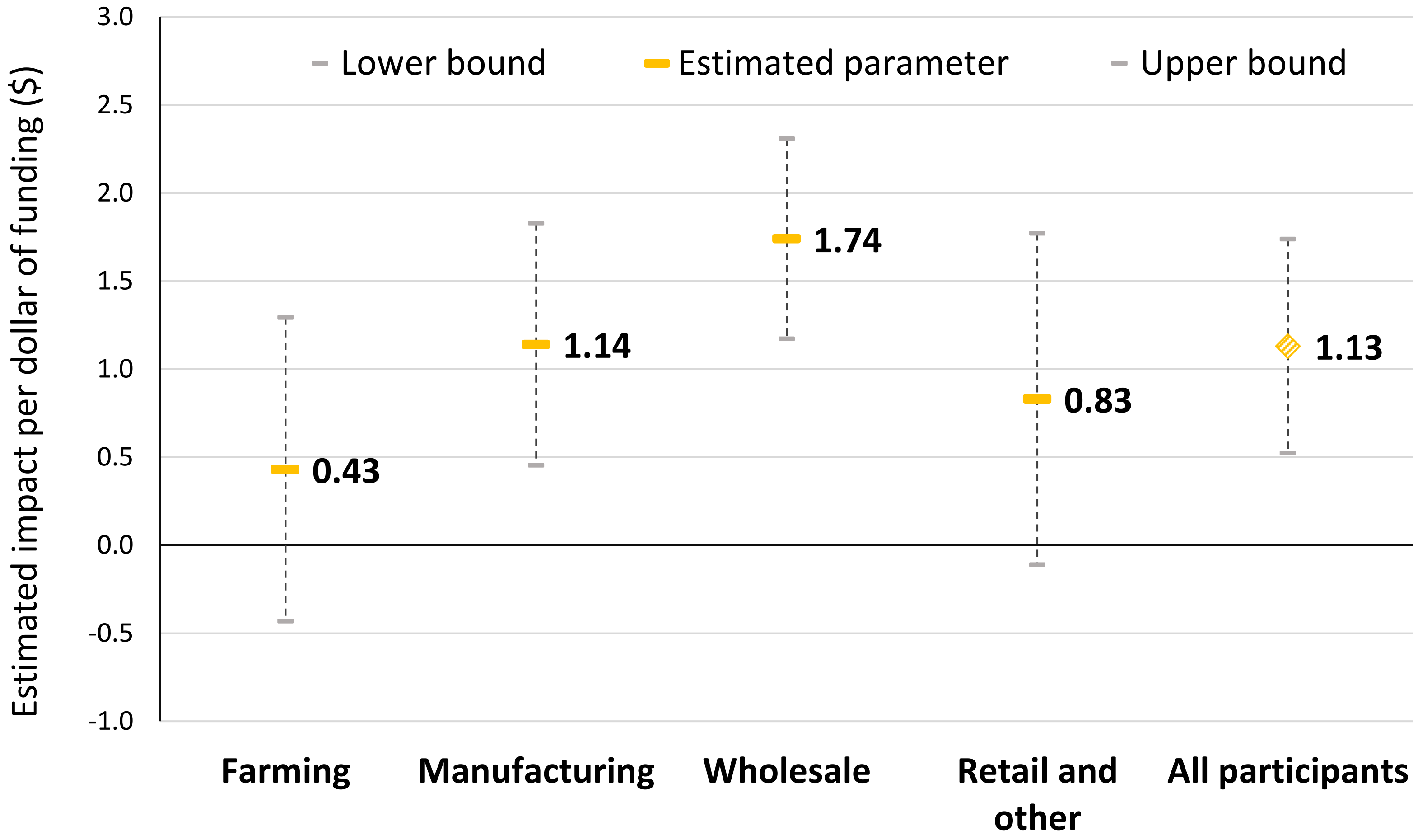
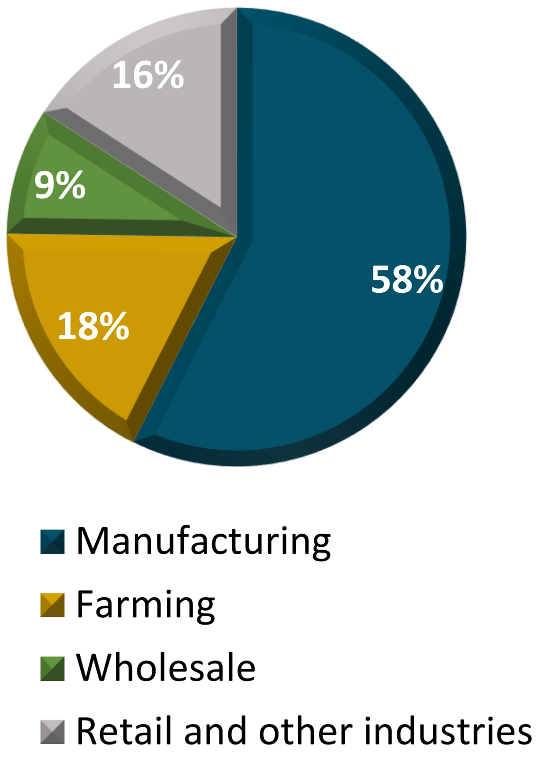


Figure 3. Share of participants by industry



Drivers: Firm size

Finding 3: The largest firms (i.e., revenues of $25 million or more) realise the largest benefits from the program.

Figure 4. Estimated program impact on an average participant's gross profit: firm size classed revenues

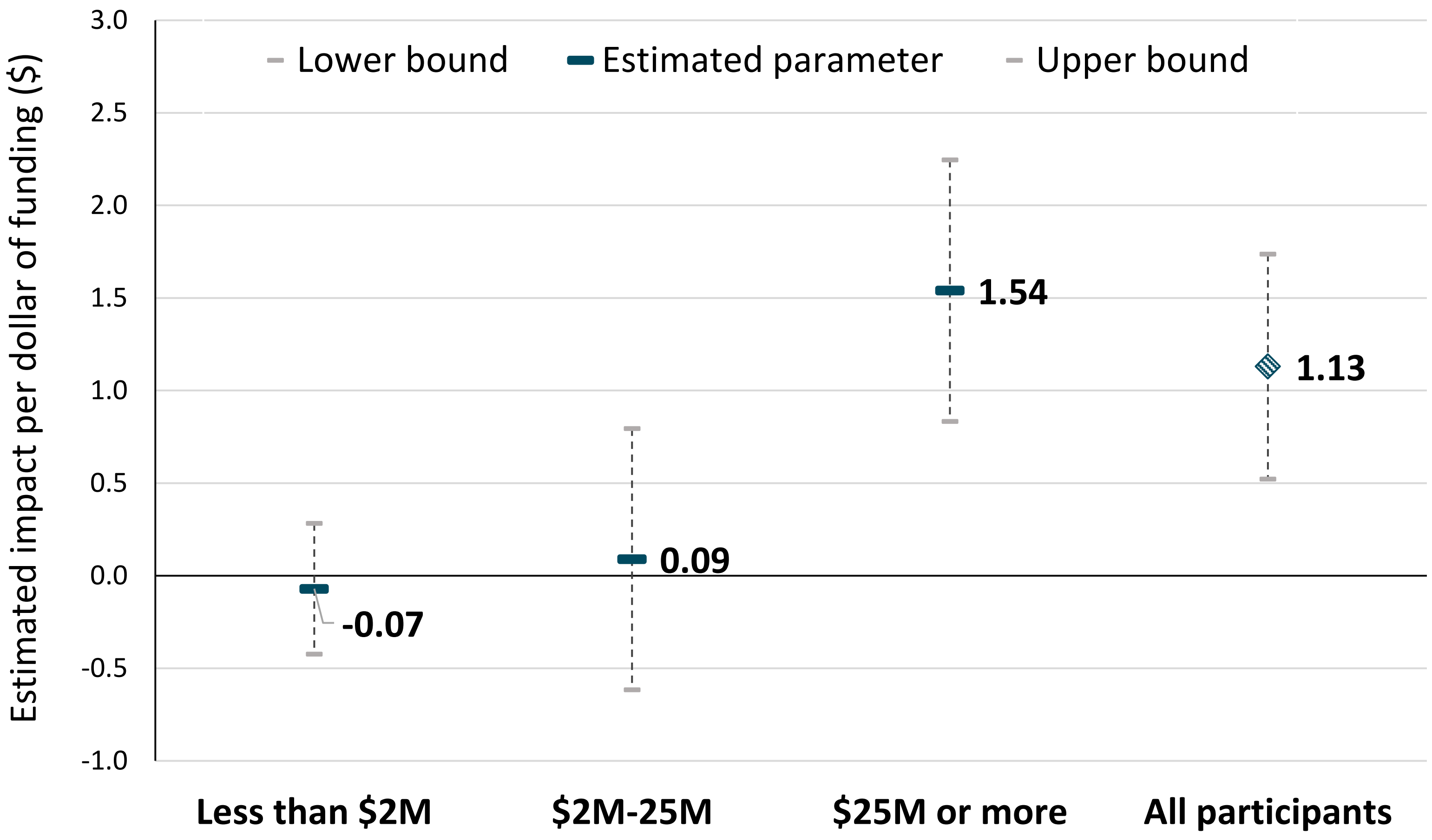
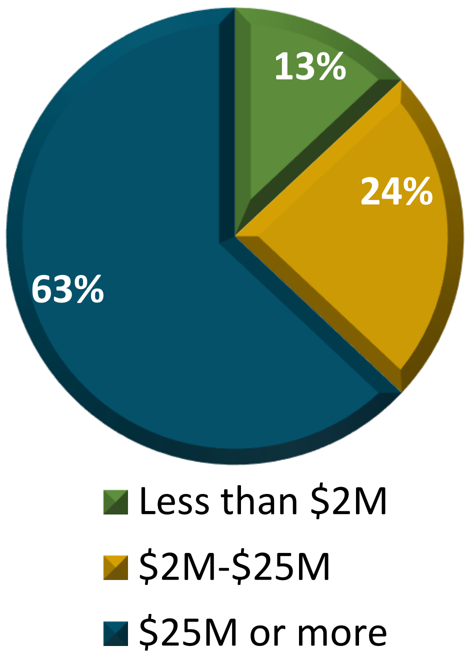


Figure 5. Share of funding by firm size



Drivers: Gender of ownership

Finding 4: Enterprises majority owned by women realise larger benefits than those majority owned by men.

Figure 6. Estimated program impact on an average participant's gross profit: Gender effects

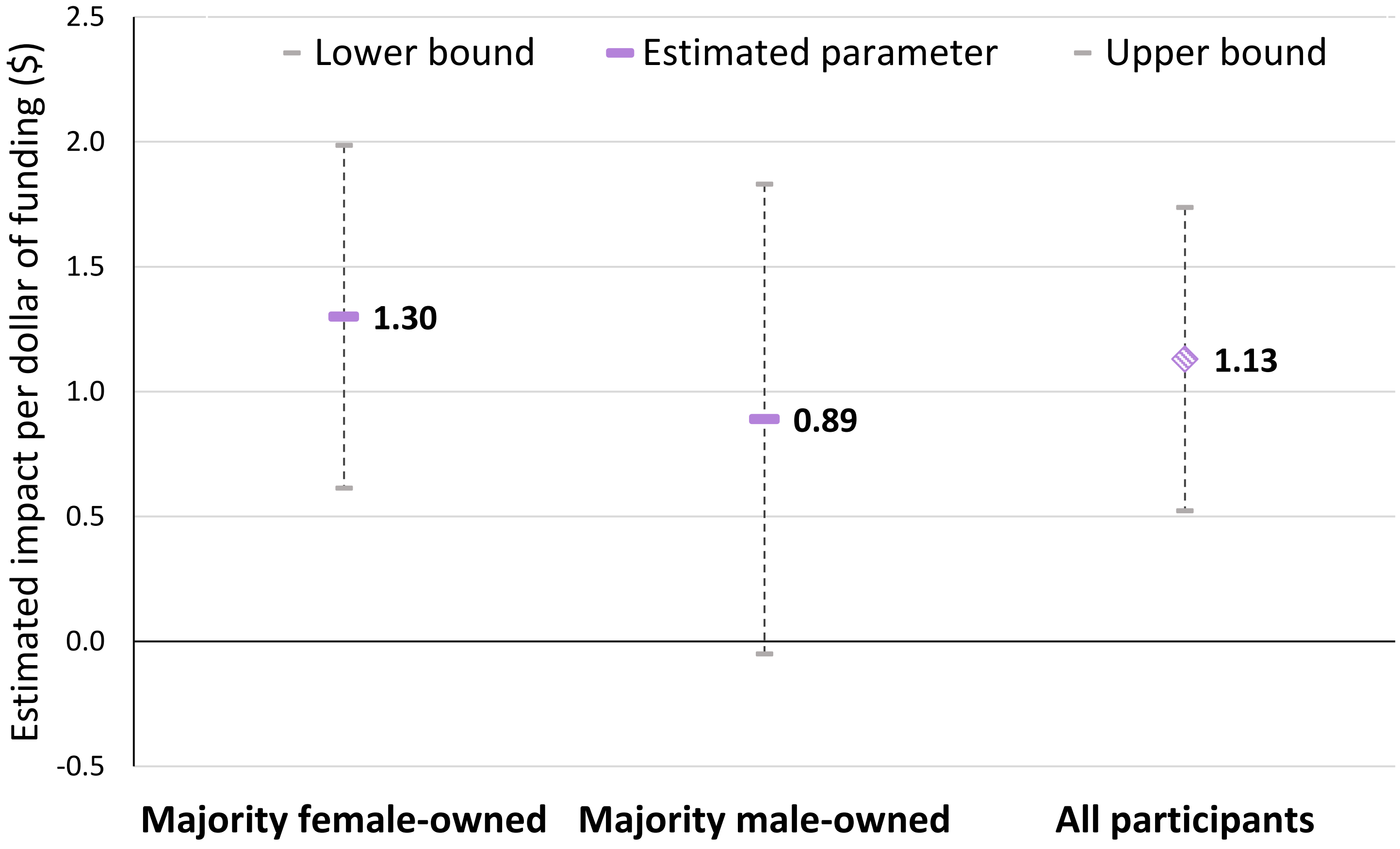
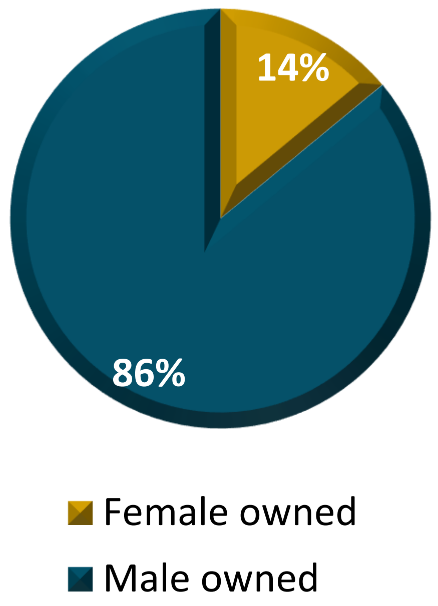


Figure 7. Share of participants by gender of ownership



Robustness tests

Several sets of tests are performed to check robustness:

Sensitivity to matching criteria

Model specification

1. Logit model for matching
2. TWFE Panel DID

Not-yet-treated as control
(see Annex 1)

Like the treated, the not-yet-treated are from the pool of winners.

Rejected applicants as treatment
(see Annex 2)

If the rejected applicants experience an increase in financial performance after rejection, self-selection bias becomes a strong possibility.

Treatment effect heterogeneity bias

Treatment effect heterogeneity bias

Problem

The DID estimator is a weighted average of all possible two-group/ two-period DID estimators in the data.

Implication

A potential bias in **staggered** designs when treatment effect varies over time and/or across cohorts of participants.

Types

1. **Over time:** when already-treated units are used as controls, and their treatment effects vary over time.

   → **Solution**: do not use the already-treated as control

2. **Across units**: cohorts treated at different times experience different treatment effects

   → **Solution**:

   i. if there is little variation in treatment timing,

   ii. if the untreated group is very large (relative to the early-treated or the late treated), or

   iii. if some cohorts are very large (relative to others).

ii. means the bias could be avoided by ensuring that the control group consists of **only the never-treated**, which is the case in this study.

Note: see **Handout 2** for more details. *Sources:* Goodman-Bacon, 2021; Baker et al., 2022; Callaway and Sant'Anna, 2021; Sun and Abraham, 2021; de Chaisemartin and D'Haultfoeuille, 2020.

Heterogeneity of treatment effect (HTE) across units

Finding 5: **Program impact differs across cohorts of participants and is still zero for the late participants.**

Figure 8. Estimated program impact on an average participant's gross profit: various cohorts

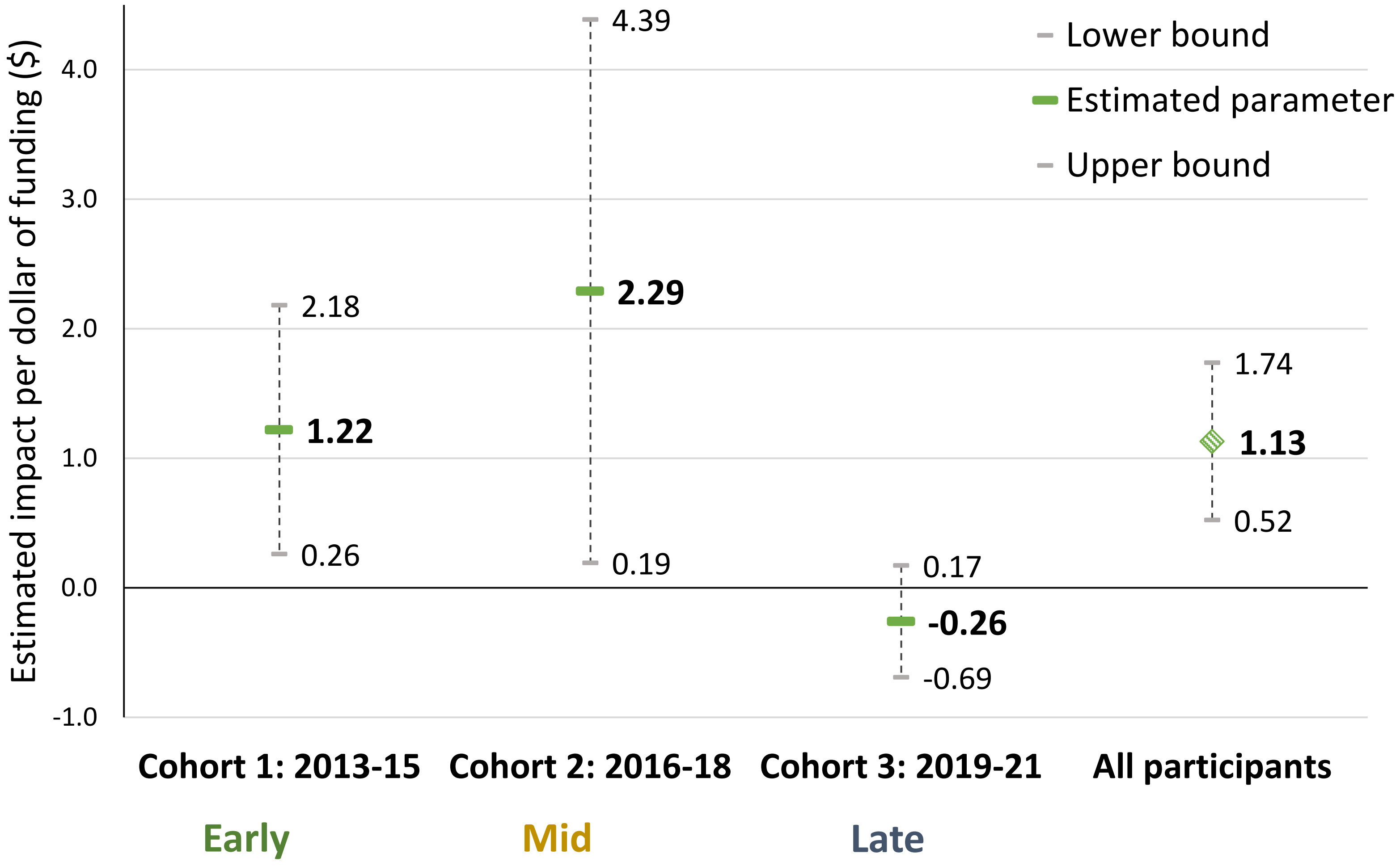
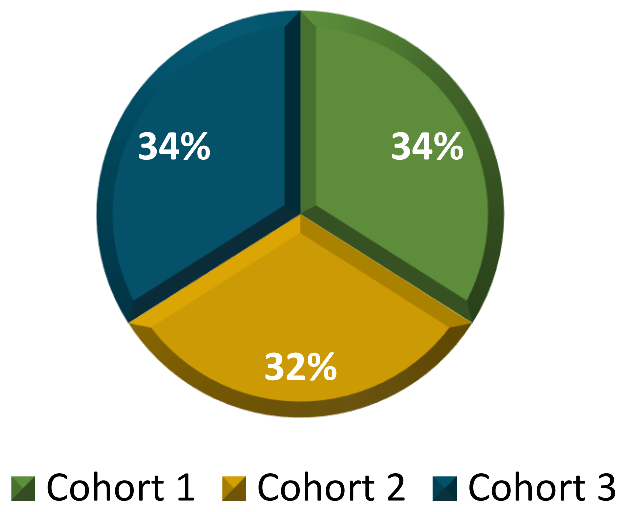


Figure 9. Share of participants by cohort



Weighted average of program impacts for 3 cohorts:
1.22*34%+2.29*32%+0*34%=**1.14**
→ No evidence of HTE bias

Summary and conclusion

	Research Questions	Answer
Phase I	*Are AAFC's AgriInnovation Stream C and AgriInnovate Programs effective in improving the economic performance of the recipient firms?*	**Yes.**
Phase II	*1. How long does it take for the program to show its impacts?*	It takes **3-4 years** for the program to have a meaningful impact on firm performance.
	2. How long does the impact of the program last?	The aggregate effect grows for **at least 7 years**.
	3. What industry is the main driver of the impact?	Wholesale > **Manufacturing** (58% of participants) > Retail and other industries > Farming.
	4. What firm size is the main driver of the impact?	Those with total revenues of **$25 million or more**.
	5. Which gender experiences a larger impact?	**Female-owned** enterprises.

Conclusion: In the evaluation of innovation programs, the **granularity of program impact** adds insights valuable to the process of evidence-based policy development.

Thank you.

Mohammad Torshizi & Jennifer Syme

Economists
Farm Economic Analysis Division
Agriculture and Agri-Food Canada
Government of Canada

Handout 1: The literature on matching difference in difference: theory, application, and comparison

Handout 2: Treatment effect heterogeneity bias

Annex 1: Robustness tests with not-yet-treated

Not-yet-treated (NYT) as control

Create a control group comprised of future *participants* in the years before they become participants (i.e., the not-yet-treated). In this fashion, both the treated and the untreated are from the pool of “winners.” Thus, they are likely to be similar with respect to their intrinsic abilities. This approach exploits the time lag in participation to address the “good managers” versus “poor managers” problem or similar patterns resulting in self-selection bias (see Table A1).

Table A1. Using the not-yet-treated in the control group

Year	**Participating firms**	**Matching**
t-1 (the year before the program)	-	-
***t* (1st year of the program)**	**Firm *A***	**Treatment group: Firm *A* Control group: Firm *B*, Firm *C***
t+1 (2nd year of the program)	Firm *B*	Treatment group: Firm *B* Control group: Firm *C*
t+2 (3rd year of the program)	Firm *C*	-

Results: The program has a statistically significant impact on the financial outcome of the participants (see Table A2).

Table A2. Results of 5 MDID regression models with 5 groups of not-yet-treated participants

	Estimated impact on total revenue	Std. Err	P-value	# of treated firms	# of NYT firms
NYT as of 2016	9.2	4.9	0.067	26	43
NYT as of 2017	**10.0**	**5.1**	**0.056**	**38**	**31**
NYT as of 2018	10.8	5.4	0.049	43	26
NYT as of 2019	11.8	4.5	0.012	49	20
NYT as of 2020	11.0	4.5	0.017	56	13

Annex 2: Robustness tests with rejected applicants

Rejected applicants as treatment

Test whether the rejected applicants experienced a significant increase in their performance after their rejection date and relative to a **control group**.

- **The control group:** firms that are like the rejected applicants with respect to observables but never applied for the program.
- **The logic**: if the increase in performance is all due to unobservable factors – e.g., having an innovative idea or having an opportunistic manager – and not affected by the program at all, then the rejected applicants would be as likely as the accepted applicants to experience an increase in their performance.
  - If the rejected applicants do experience a significant increase in their financial performance after rejection date and compared to a control group, self-selection bias becomes a strong possibility; the larger the magnitude of the increase, the larger the size of the bias.
  - However, the rejected applicants are often not as good as the accepted ones in one (often observable) way or another – that's why they are rejected. Thus, if the rejected applicants do not experience an increase in their performance, it could be attributed to the reason for which they were rejected.
- **Regression results**: no statistically significant effect for the rejected.