CO Women's Chamber of Commerce - CU Denver Statistical Consulting, Survey Design Report

Chandler Zachary*
MATH 6330
CU Denver

Submitted May 15, 2020

*I am grateful to Kristen Blessman and Deb Siverson for your hard work. We truly were a team. Special thanks also to Prof. Stephanie Santorico for helping me maintain focus and work ethic in a tumultuous time.

1 Introduction

The Colorado Women's Chamber of Commerce (CWCC) is a professional organization that advocates for women in business across Colorado.[1] Among their top goals are to propel women into the highest leadership roles of Colorado companies and accelerate the growth of women-owned business in the state. CWCC cites several statistics on its website in support of these goals:

- In Colorado, 86% of women-owned businesses have only one employee.
- Women make up only 14% of CEOs.
- Colorado is nearly the last in the nation for the number of women on publicly traded boards.
- Women influence 85% of the purchasing decisions and women-owned businesses generate $4.7 billion in revenue ...

The national US Women's Chamber of Commerce website cites articles from 2016[2] that offer some corroboration of these statistics. Namely, women-owned businesses largely have a more difficult time than men-owned businesses in growing annual revenue to $1,000,000, women-owned businesses struggle with access to capital financing, and only 10% of women-owned firms have employees. To ameliorate these conditions, the CWCC hosts a variety of engagements and leadership training programs designed to foster a community whose mission is to advance women in business.

Members can enter the organization through a variety of membership levels, segmented into small business and corporate categories. CWCC programming consists of several channels of engagement, such as speaking events, peer groups, and executive leadership training. In January of this year, two members of CWCC leadership approached UC Denver students

[1] https://www.cwcc.org/
[2] https://uswcc.org/uswcc-in-the-news/

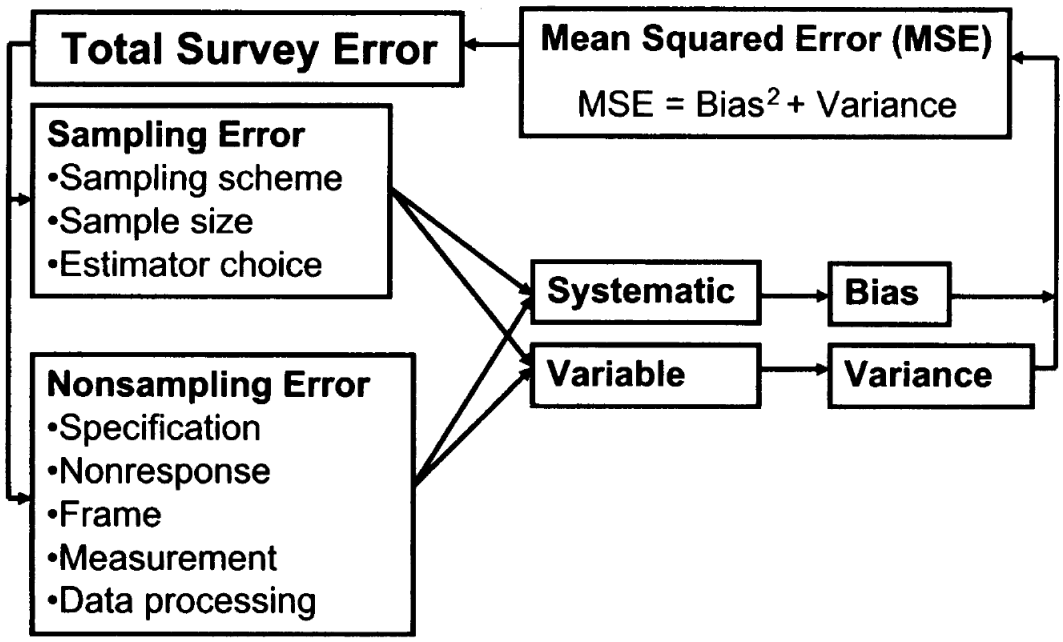
to ask for help in measuring their progress at delivering on their mission and the quality of its programming. In previous years, CWCC had given surveys to its membership, but these surveys lacked quantitative and objective measures of the impact of the programming on women's business and career outcomes. CWCC leadership wanted to completely redesign their surveys to take advantage of the rich, untapped data they felt their members could offer, and they asked CU Denver students to supply the tools to acquire those data.

I was invited to work together with the CWCC leadership to redesign the survey process and develop a framework for collecting data. My role was to provide input on updating the questionnaire to capture relevant, objective business and career measures and conduct an analysis of the survey data to help leadership assess the efficacy of their programming. In February, the design team held its first meeting, consisting of Kristen Blessman, President and CEO of CWCC, Deb Siverson, founder and CEO of member business Xponents, Inc., and myself. During this meeting, I gathered information on the needs of the survey, the goals of CWCC leadership, and the difficulties anticipated in collecting survey information. After two initial meetings, I began a draft of survey questions that served as the palette for the new survey. We have since held several meetings to edit and revise questions.

While we had an aggressive timeline to finish the survey and process data by May, the spread of COVID-19 had a detrimental effect on the business community of Colorado. We had to shelve our plans while Kristen and Deb focused on more immediate needs of supporting members and helping businesses navigate through the abrupt change. Our revised timeline is in flux, but we remain committed to implementation sometime this summer. I have begun collecting pre-testing feedback and hope to continue this fruitful process.

The following report outlines my approach for the survey design. I closely follow Biemer and Lyberg's *total survey error* framework throughout survey construction (Biemer and Lyberg, 2003). I have found this to be an indispensable tool for organizing information to connect errors in survey and questionnaire design with fundamental statistical principles. Figure 1 depicts a summary graphic for this framework. The overall report has the following

Figure 1



Source: Total Survey Error: Design, Implementation, and Evaluation. Biemer, Paul P. (2010).

structure. This section introduces the goals of the CWCC-CU Denver Statistical Consulting relationship and summarizes our progress so far. Section 2 outlines the various components of the survey design with some exposition where necessary. Section 3 lists potential sources of non-sampling error in the survey, and Section 4 summarizes our approach to sampling design. Section 5 closes.

2 Survey Design

Survey Objective

For the 2020 year, the CWCC Board identified two key performance measures for evaluating the efficacy of the CWCC programming. These measures are:

- Help women-owned businesses reach $1,000,000 or more in annual revenue, and
- Help women enter senior leadership echelons in their industries.

The design team endeavored to determine the factors in CWCC programming that drive those outcomes. The questionnaire is tailored to each category of membership, small business and corporate. The purpose behind this delineation is to measure variables that may be unique to each business environment. Needs of each member type may overlap but should not be considered identical or uniform. CWCC leadership can use this design to tailor resource deployment to the needs of each member.

Population of Interest

In ideal conditions, the population of interest would consist of Colorado women at any working age, level of education, marital status, family size, race, and ethnicity, involving both members non-members of CWCC. This would enable a random sample of working women for comparison of groups who are the same along all dimensions except whether they are members. This is the usual design implemented when inference is the analytical goal. There are practical considerations for obtaining a random sample from this population. I discuss this at greater length in the Sampling Design section below. To accommodate these consideration, we have narrowed the population to CWCC membership.

Mode of Collection

The survey will be conducted via email. We will distribute the survey to CWCC members using Survey Monkey with the most recent email address list. Survey data will be received, cleaned, and saved as comma separated values files. We have not yet determined which software package will be used to process and analyze the data.

Measurement Variables

There are four categories of measurement variables: outcome, predictor, control, and demographic. Outcome variables measure quantities directly related to the Board's key per-

formance measures. Specifically, survey questions ask respondents about their revenue and growth in revenue and whether they received promotions. These questions directly measure whether business are earning over $1,000,000 and being promoted into leadership roles. For small business owners, the first question asks:

> What was your total revenue in your most recent fiscal year?
>
> - $< \$250k$
> - $\$250k - 500k$
> - $\$500k - 1MM$
> - $> \$1MM$

By structuring the responses this way, we can easily create a binary indicator signifying whether the respondent earned over $1,000,000.

The promotional questions use categories of job titles taken from the study "Women in the Workplace" (McKinsey, 2019) with one extra category included by the design team to capture promotion at the lower levels of leadership. This was included to account for differences in job titles among companies and to measure the "bottleneck effect" exhibited by women at the lower stages of leadership. McKinsey (2019) shows that women seldom achieve senior levels of leadership because there is little advancement beyond lower levels. We wanted to identify whether this is happening among CWCC members. Specifically, the promotion question asks:

> If you were promoted, to which level were you promoted?
>
> - Team Leader/Supervisor
> - Manager
> - Director
> - Vice President
> - Senior Leader

- C-suite

It is important to note that this specific question is conditioned on two prior questions that ask whether the respondent was even considered for a promotion and, if so, whether she received it. There is a key advantage to this question ordering beyond directly measuring the Board's target: we can directly assess whether and to what extent members are being promoted at all. This increases our accuracy at evaluating the deployment of CWCC resources for developing leadership and executive qualities among members that translate into promotion.

Developing quantitative predictor variables to measure CWCC programming was mission critical in survey design. Kristen shared with us that their prior surveys were inadequate for capturing drivers of business success because they failed to connect business outcomes with the training and speaking events hosted by CWCC in an objective way. This provided the design team a distinct opportunity to ask a range of questions of the membership that yield some basic descriptive information, as well as quantitative assessments of events and training. We included questions about the respondent's type of CWCC membership, attendance at events, and career goals and priorities.

Following this sequence of questions, we intend to ask the respondent to provide a measure of how well the CWCC programming has helped them achieve their goals. There are currently three forms of this question that we intend to get feedback on during pretesting. Each question has some distinguishing element intended to elicit a response that directly connects the respondent's career progress with her experience in CWCC. The first question type simply asks the respondent to choose from three broad attitude or belief statements. The second two questions have numeric responses, but they differ in how they connect to the CWCC programming. In particular, they distinguish, between programming and engagement to provoke in the respondent a sense of magnitude of her connection to events and training. We want to capture whether there is a difference for members between simply attending events and the types of events and frequency of attendance. Following this

question, we ask respondents to choose the specific skills CWCC has helped them develop and use a free-response question to tie a specific event in their career to something they learned via CWCC programming. This range of questions should provide a rich resource for examining the factors that contribute to the Board's outcomes of interest.

Demographic questions are designed to add descriptive elements to the survey to contribute additional information resources to CWCC. Business attribute questions supplement demographic information so that CWCC leadership have a better picture of who their members are, the types of businesses and industries in which they are employed, and the variety of products and services represented. This is an improvement not only for CWCC, but for the business community of Denver. My research revealed no organization or agency that keeps this information about the business community. This could help CWCC become a resource for other business advocacy groups in the area.

Control variables are included to account for sources of variation that may be correlated with variables that measure advancement and revenue growth. For example, our questionnaire asks respondents to provide their role or department and their key performance assessment measures. The type of work and the business or operational risk associated with a respondent's work can be reflected in their compensation and career advancement because the labor market and employer may place a higher value on some skills and responsibilities relative to others. To the extent this is happening among CWCC members, we want to prohibit those sources of explanatory variation from contaminating the measurements of interest. A table of the full set of current variables, their definitions, data types, and membership affiliations is provided in Appendix Table A1.

Pre-Testing

Prior to the end of the semester, the design team pre-tested the questionnaire with a small-business member and a corporate member. The questionnaires were released via email, and I personally spoke with both respondents about their experience. The small-business mem-

ber identified questions that were intended for both small business and corporate members but whose responses failed to include characteristics of small business owners. This led to separating the relevant questions, tailoring them to the appropriate audience. The corporate member gave feedback at length about the goals of the CWCC programming in developing "executive presence" among members and about eliciting a level of engagement. Executive presence refers to the sense of confidence and attitude when expressing views to an audience to affect change, especially among people who may disagree or otherwise need persuading. This also refers to leadership training designed to help members learn how to navigate change and set long-term strategy for businesses. In her feedback, the corporate member shared that she believes these characteristic will help close the gap between men and women representation in senior leadership. The level of engagement refers to members who have varying degrees of involvement with CWCC, such as attending only speaking events or a few events per year versus committing to the various training courses offered by CWCC intended to develop skills over time. This feedback prompted the addition of questions to address these specific items of inquiry. We are currently waiting to pre-test the updated questionnaire among other members before releasing the survey to the full membership.

3 Sources of Non-Sampling Error

Non-sampling error refers to sources of error in the measurement of the quantities of interest and can have effects on the bias and variance of estimators used for inference in a survey. The following section provides my assessment of the risks posed by non-sampling error sources so far. For reference, "systemic" error can be connected with bias, which refers to whether survey questions on average over- or underestimate the quantity of interest. "Variable" error can be connected with variance and refers to the breadth of variation in item responses.

Specification Error

The risk of specification error is potentially high for the systemic component and low for the variable component. This is a pioneering effort for CWCC, so there will likely be a learning curve while CWCC determines the relevant measures and establishes a baseline for measuring itself. The survey *outcome* variables are straightforward because the Board has established clear definitions for gauging their performance, and they can be easily measured with binary indicators in the data. By contrast, matching these outcomes to their underlying drivers is unclear. Members are likely to differentially assess their perceived challenges to growth and progression and the role of CWCC in aiding growth and progression. It is possible there are important factors we have not even considered and have thus omitted. The design team has attempted to mitigate specification error by asking a range of questions about respondents' interaction with CWCC and underlying business concerns. We hope this will offer several ways for respondents to provide insight and establish common definitions. Until we have more pre-teseting feedback, I assign a medium level of systemic risk and a low level of variable risk to specification error.

Measurement Error

For similar reasons to those listed above, the risk of measurement error is potentially high. I believe there is little threat from cognitive bias in many questions because of their factual, objective nature. To the greatest extent possible, we utilize a variety of word choice and item response types overall. Some questions are necessarily challenging because they ask respondents to translate subjective, time-variant experiences into objective item responses, which can be difficult. While we wait for additional pre-testing feedback, it is difficult to assess effects from factors such as question ordering, fatigue, and language. For now, I am assigning a high risk score to measurement error and will revisit this after pre-testing is complete.

Frame Error

Lack of a suitable control group motivates a non-random sample and the use of a prediction model. This is further detailed in the Sampling Design section. The choice of a prediction model eliminates the systemic risk from frame error because estimators will be mechanically biased, but variable error may be high if the survey does not sufficiently cover all parts of the CWCC membership.

Non-Response Error

CWCC leadership has emphasized the threat posed by non-response. Members consider some business information, such as revenue, to be strategically valuable and are reluctant to disclose it. Furthermore, some members feel their reputations could be damaged by disclosing too much information. There is also a general concern of personally identifying information being hacked. These combined factors can have a harmful effect on response rates. The design team has taken two measures to mitigate the risk of non-response. First, no personally identifying information is collected in the survey, either of the respondent or her employer. This should affect identity concerns. Second, many item responses have been deliberately designed to be categorical. Since respondents need not give absolute figures, we think this will provide an additional later of anonymity. It has also been recommended that we address survey emails specifically to individuals by using their names[3] and conduct follow-ups for non-responses. In the absence of the effects on the business climate of COVID-19, I would gauge these efforts to be effective at minimizing non-response risk. However, I remain concerned that some members will still be feeling the effects of the pandemic and may not prioritize the survey. Therefore, I assign a medium score to systemic risk and a low score to variable risk.

[3]Names will not be collected as part of the survey results.

Data Processing Error

For now, evaluating the scope of risk from data processing is unclear. The complexity of the survey has increased since the beginning of the project. We have chosen to add multiple selection options to some item responses. Without a technical expert managing the data cleaning and processing function, there is a potentially high risk of error. I am currently examining different software packages for processing the data and conducting analysis. CWCC leadership is familiar with Survey Monkey for collecting survey data and Excel for cleaning. However, in my absence, Kristen and Deb will lack a subject matter expert for making informed decisions about the data that can affect the quality of the results. I assign a high risk score to this source of error and will revisit this as more information become available.

Table 1 gives a summary of the sources of non-sampling error and the risks posed to systemic and variable error. For reference, the four risk scores are, in order: None, Low, Medium, High.

Table 1

Survey Error Component	Systemic Risk	Variable Risk
Specification	Medium	Low
Measurement	High	High
Frame	None	Low
Non-response	Medium	Low
Data Processing	High	High

Survey Error Component Risks

4 Sampling Design

The purpose of random sampling in any survey design is two-fold. First, it ensures that the sample is representative of the population. Second, it ensures that unobservable characteristics that might be correlated with the measurements of interest are controlled for, preventing contamination of inference. These two things are accomplished because random design elim-

inates the possibility of some characteristics from being over- or under-represented in the sample, which prevents selection bias, the tendency for some people to be surveyed or not in some correlated way. At least in theory, this should be true. In practice, random design is fragile. As stated previously, the ideal sample for the CWCC survey would have been a group of women from Colorado, randomly chosen, who are alike in every way except that some are members of CWCC, and some are not.

Feasibility constraints prevent such a design. There are county-level data available through the American Community Survey (ACS) and Current Population Survey (CPS) on attributes such as age, income, occupation, and family size that could facilitate sampling along enough dimensions to compare members with non-members. However, we have no contact information for non-members to ask them to take the survey. Without such information, we are left with a convenience sample, a non-random sample of members who are accessible because we know how to contact them. The consequences of a non-random sample are biased estimates of the effects of CWCC membership and unreliable standard errors in linear regression.

Mathematical techniques exist for calibrating this sample to reflect the demographic and business attributes of the overall state of Colorado. However, we still lack information for non-members, an important control group if we are to determine the *effect* of CWCC membership. To sidestep this issue, it was proposed that we use a prediction methodology rather than an treatment-control inferential comparison. This is advantageous in several ways. First, prediction models are uncontaminated by a non-random sample because the goal of prediction is to minimize prediction error, not supply estimators of population parameters. Thus, we can make the best use of the sample of members without finding non-members to form a control group. Second, prediction may actually serve the purpose of determining the factors of revenue growth and career advancement better than inference. This is because prediction requires finding the relevant predictor variables and functional form that produce the best accuracy in when supervised by the outcome rather than producing the best es-

timators (i.e. - hypothesis tests and confidence intervals). This puts the responsibility on the design team to specify and measure variables correctly, and the model can tell us which variables are useful in predicting the outcome. A third advantage is that prediction models are easier to interpret for people without a rigorous statistical background, making it easier to maintain the model in my absence.

At this stage, we have not yet decided on a particular prediction model. Table 2 updates the sources of survey error to include prediction effects on sampling error. I evaluate the systemic and variable risks to be low.

Table 2

Survey Error Component	Systemic Risk	Variable Risk
Specification	Medium	Low
Measurement	High	High
Frame	None	Low
Non-response	Medium	Low
Data Processing	High	High
Sampling Error	Low	Low

Survey Error Component Risks, Including Sampling Error

5 Closing Remarks

This document is intended to be a summary of the current progress toward achieving the survey objectives and to document the justification for decisions made in questionnaire design and construction. It is not final because the survey has not been finalized. I intend to provide updates as new information becomes available.

It has been my genuine privilege to work together with Kristen and Deb. I have found this project fruitful for solidifying my foundation in important statistical concepts, as well as preparing me to execute a professional, non-academic project.

References

Biemer, Paul P. (2010). Total Survey Error: Design, Implementation, and Evaluation. *The Public Opinion Quarterly, 74*(5), p817-848. https://doi.org/10.1093/poq/nfq058.

Biemer, Paul P., and Lars E. Lyberg. *Introduction to Survey Quality.* Hoboken, John Wiley & Sons, Inc., 2003.

Buskirk, T. D., Kirchner, A., Eck, A., and Signorino, C.S. (2018). "An Introduction to Machine Learning Methods for Survey Researchers." *Survey Practice, 11*(1). https://doi.org/10.29115/SP-2018-0004.

Appendix

Variable	Definition	Data Type	Variable Type	Member Type
Revenue	Respondent's revenue in most recent fiscal year	Categorical	Outcome	Small Business
RevGrowth	Growth in percent of respondent's revenue in most recent fiscal year	Continuous on [0, 1]	Outcome	Small Business
EmpGrowth	Employees added to respondent's business in most recent fiscal year	Categorical	Outcome	Small Business
ConsProm	Whether respondent was considered for a promotion in previous two years	Binary	Outcome	Corporate
Promote	Conditioned on *ConsProm*, whether respondent was promoted	Binary	Outcome	Corporate
LevProm	Conditioned on *Promote*, to which level was respondent promoted	Categorical	Outcome	Corporate
MemberTerm	Length of time respondent has been a CWCC member	Continuous	Predictor	Both
AttLastYear	Whether respondent attended a CWCC event in previous year	Binary	Predictor	Both
NumAttLastYear	Conditioned on *AttLastYear*, number of CWCC events attended last year	Integer	Predictor	Both
WhichAttLastYear	Conditioned on *AttLastYear*, types of CWCC events attended last year	Categorical	Predictor	Both
Goal	Respondent's top business or career goal	Categorical	Predictor	Both
JoinReason	Respondent's reason for joining CWCC	Categorical	Predictor	Both

Table A1: Potential Variables and Definitions

Variable	Definition	Data Type	Variable Type	Member Type
OtherMem	Whether there are other CWCC members at respondent's workplace	Binary	Predictor	Both
Attitude	Respondent's attitude about CWCC helping to achieve goals	Categorical	Predictor	Both
ProgGauge	Respondent's gauge of CWCC's help in progress toward goals	Categorical	Predictor	Both
PercSuccess	Respondent's percent of success attributed to CWCC programs	Categorical	Predictor	Both
Skills	Skills CWCC has helped respondent improve	Categorical	Predictor	Both
Role	Respondent's department or role at company	Categorical	Control	Corporate
KeyPerfMeas	Respondent's key performance measures	Categorical	Control	Corporate
Industry	Respondent's industry	Categorical	Control	Corporate
Field	Respondent's field	Categorical	Control	Small Business
Race	Respondent's self-identified race	Categorical	Demographic	Both
Ethnicity	Respondent's self-identified Hispanic ethnicity	Binary	Demographic	Both
Age	Respondent's age group	Categorical	Demographic	Both
Education	Respondent's highest education level	Categorical	Demographic	Both
CompEmp	Number of employees at respondent's company	Categorical	Control	Both
ProdServe	Whether respondent's company sells a product or service	Binary	Control	Both

Table A1: Potential Variables and Definitions, Continued