Douglas R. Hoffer ❖ Vermont State Auditor

Measuring Up:

Examining performance reporting in the Vermont Annual Outcomes Report

24 January 2022

❖

Investigative Report 22-01

Mission Statement: The mission of the Auditor's Office is to hold State government accountable by evaluating whether taxpayer funds are being used effectively and identifying strategies to eliminate waste, fraud, and abuse.

Investigative Report: An investigative report is a tool used to inform citizens, policymakers, and State agencies about issues that merit attention. It is not an audit and is not conducted under generally accepted government auditing standards. Unlike an audit, which contains formal recommendations, investigative reports include information and possible risk-mitigation strategies relevant to the topic that is the object of the inquiry.

Principal Investigator: Fran Hodgins, Government Research Analyst

Table of Contents

A Note from the Auditor

Performance measurement – when done well – is a powerful tool that leads to evidence-based decision making, better program management, and greater accountability. When poorly executed, however, performance measurement can be a waste of time, leaving users with incomplete, unhelpful, or (at worst) misleading information.

Vermont State government produces a dizzying number of performance-related documents each year, including the Annual Outcomes Report. Per Act 186 of 2014, the Outcomes Report is intended to provide the Vermont legislature with data to "know how well State government is working" to achieve ten population-level outcomes (3 V.S.A. § 2311). Population-level analyses look at outcomes for a whole population (e.g., all Vermonters or young Vermonters) rather than looking at outcomes from a specific government program or policy.

Now in its eighth year, **we examined the quality and usefulness of the Annual Outcomes Report** using nine principles of effective performance measurement. In addition to our analyses, we met with a bipartisan group of Vermont legislators to better understand the extent to which legislators themselves feel well-served by the Report. We found:

1. **INFLUENCE:** Population-level analyses have limited utility as a tool for State government performance measurement because **the State has limited influence over most population-level outcomes.**
2. **CLEAR GOALS AND OBJECTIVES:** The **population-level goals are broad and undefined**, which leaves the door open for multiple, possibly even contradictory, interpretations of success.
3. **SIGNIFICANT AND RELEVANT:** Many indicators **do not capture the most significant or relevant information.**
4. **COMPARABLE:** Most indicators **lack contextual information and comparative data** that would enable readers to draw meaningful conclusions.
5. **DISAGGREGATED:** When appropriate, **disaggregating data would inform policy discussions** about the disparate and unintended inequitable impacts of policies and programs and/or where to target resources.
6. **UNDERSTANDABLE:** There are **opportunities to improve how data is presented** to make it easier for the reader to understand.
7. **WELL-DEFINED:** Some indicators have **no definitions of the variables being measured**, without which the reader is left guessing.
8. **VALID:** Many indicators **lack specific data sources or methodology information**, making it difficult for the reader to assess the validity of the data.
9. **TIMELY:** The data is generally current, though **some descriptions and indicators need updating.**

As one legislator we spoke with said, **it's time to "go back and scrub [the Outcomes Report]. Ask whether we are getting what we want from these indicators."**

Introduction

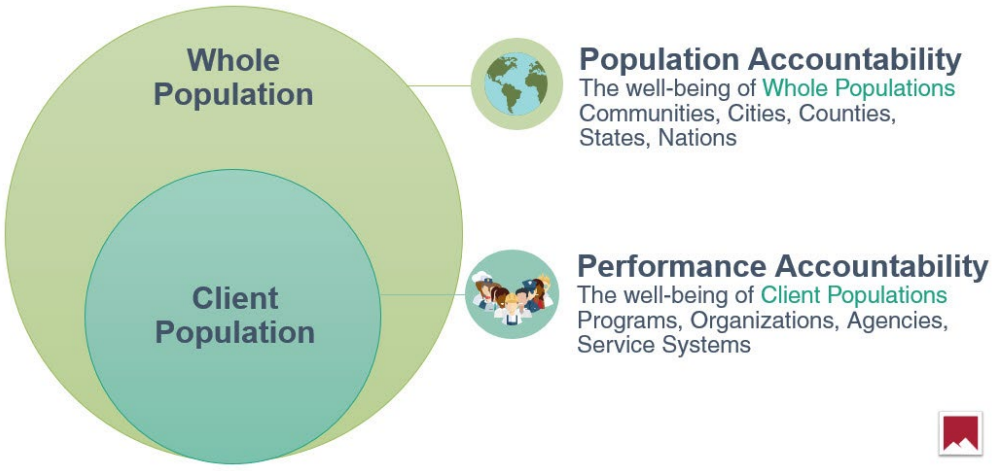
In the context of State government, performance data can be used by legislators to design policies and programs, assess whether existing programs are working as intended, and make decisions about resource allocation. At the agency level, program staff and leadership can use performance data to inform strategy implementation and continuous improvement efforts. Performance measurement is also central to the work of the State Auditor's Office. Having access to accurate, relevant, and timely data helps our office hold State government accountable and assess whether taxpayer funds are being used effectively.

This report examines the quality and usefulness of the Annual Outcomes Report, which is produced by the Agency of Administration's Chief Performance Officer using outcome indicators selected by the joint legislative Government Accountability Committee. In addition to our own analyses, we met with a bipartisan group of Vermont legislators in one-on-one sessions to better understand the extent to which legislators themselves feel well-served by the current outcome indicators.

Performance measurement in State government

Over the last decade, there has been a movement to improve how Vermont State government conceptualizes and measures government performance. The Legislature and Agency of Administration selected the Results Based Accountability™ (RBA) framework to guide these efforts. Like other performance measurement frameworks, RBA starts with identifying the desired end result (or outcome) and then identifying "indicators" to measure progress towards that goal. Per the RBA website, though, "what separates RBA from all other frameworks" is the use of two levels of analyses: population accountability and performance accountability.

Figure 1. RBA uses two levels of analysis: population accountability and performance accountability



Source: Clear Impact, "What is Results Based Accountability?" Accessed December 2021.

The focus on population accountability is central to the RBA framework. In Vermont, Act 186 of 2014 codified ten desired population-level outcomes, such as "Vermonters are healthy" and "Vermont is a safe place to live" (3 V.S.A. § 2311). The Government Accountability Committee is charged with maintaining a list of corresponding population-level indicators that measure progress towards the outcomes (2 V.S.A § 970). For example, a population-level indicator for "Vermonters are Healthy" is the percentage of adults who smoke cigarettes. The rate of violent crime is a population-level indicator for "Vermont is a safe place to live." These outcomes and indicators are compiled into the Annual Outcomes Report, which is prepared for the Legislature by the Chief Performance Officer in the Agency of Administration (3 V.S.A. § 2311). The Annual Outcomes Report, which is the focus of our analysis, is most easily viewed using the interactive scorecard website.

Vermont's population-level outcomes

1. Vermont has a prosperous economy.
2. Vermonters are healthy.
3. Vermont's environment is clean and sustainable.
4. Vermont is a safe place to live.
5. Vermont's families are safe, nurturing, stable, and supported.
6. Vermont's children and young people achieve their potential.
7. Vermont's elders live with dignity and in settings they prefer.
8. Vermonters with disabilities live with dignity and in settings they prefer.
9. Vermont has open, effective, and inclusive government.
10. Vermont's State infrastructure meets the needs of Vermonters, the economy, and the environment.

Measurement of agency and program-level activities predates Vermont's adoption of the RBA framework and takes many forms. In the 1990s, 32 V.S.A. § 307 required State agencies to submit a strategic plan including a description of performance measures (but no data) as part of their budget proposal.[1] Some programs have additional reporting requirements as a condition of state or federal funding. Others are not required but opt to produce annual reports or post additional data on their websites. Some agencies – like the Agency of Human Services – have adopted the RBA framework for program-level performance measurement. However, the RBA framework is not used consistently across State government for program-level reporting.

Best practices in performance measurement

Performance measurement can tell a powerful story about an organization's inputs, outputs, and outcomes. In other words, it answers basic questions about how much was spent, what it was spent on, and what we got in return. While these questions may seem straightforward, developing useful and meaningful performance data requires time, resources, and expertise (as does reviewing performance data). To assess the quality and usefulness of the Annual Outcomes Report, we drew on best practices identified by the Government Accountability Standards Board, the National State Auditors Association, and Washington State's Performance Measure Guide to develop the following criteria:

[1] The State Auditor's Office conducted two reports on this topic in 1995 and 1996.

1. **INFLUENCE:** Does the entity have the ability to influence performance in a meaningful way?
2. **CLEAR GOALS AND OBJECTIVES:** Are there specific, measurable, and well-defined goals and objectives?
3. **SIGNIFICANT AND RELEVANT:** Does the measure capture significant and relevant information that helps us understand whether the objective is achieved?
4. **COMPARABLE:** Can the data be put in context (e.g., compared to a target or to another jurisdiction)?
5. **DISAGGREGATED:** When appropriate, is data disaggregated to show how different groups are impacted?
6. **UNDERSTANDABLE:** Is the measure easy to understand? Is the data presented clearly?
7. **WELL-DEFINED:** Are the measures specific and defined?
8. **VALID:** Is the data valid? Is information about the data source and methodology provided?
9. **TIMELY:** Are the data and indicators updated regularly?

In the sections below, we use examples from the 2021 Annual Outcomes Report to illustrate our findings. Images sourced directly from the Outcomes Report have a red border like this:

Annual Outcomes Report: "Indicator Name"

[Image from 2021 Annual Outcomes Report]

Annual Outcomes Report: Summary of findings

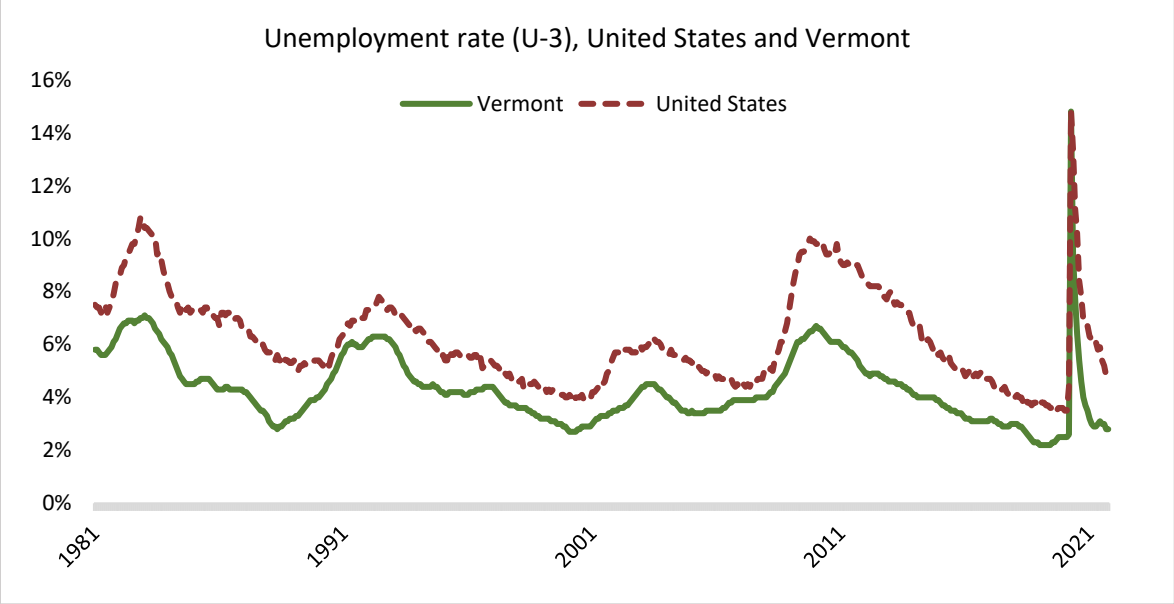
1. INFLUENCE: **The population-level outcomes and indicators have limited utility as a tool for State government accountability because the State has limited influence over most population-level outcomes.** The stated purpose of Act 186 of 2014 and the Annual Outcomes Report is to provide the General Assembly with data to "know how well State government is working to achieve the population-level outcomes." However, because population-level data are affected by many factors beyond the State's control, it is difficult to discern the impact of State activities.[2] This is not to say that State policies have no effect, nor that population-level data are useless. Population-level data *do* provide valuable information about conditions in Vermont that should inform policy discussions. But suggesting that

[2] For example, Vermont's economy is subject to many macro-level factors such as federal spending, interest rates, trade agreements, and currency exchange rates.

State-level policies are the primary driver of these outcomes could be misleading and result in misguided policymaking.

Example: The unemployment rate is one of the population-level indicators used to measure "economic prosperity." Comparing Vermont's unemployment rate to the national unemployment rate, we see that Vermont always mirrors the national unemployment rate, suggesting that Vermont's unemployment rate is primarily impacted by national trends and structural conditions rather than State-level activity. Put another way, it would not be fair to criticize the Vermont Department of Labor or the Agency of Commerce and Community Development if the unemployment rate ticks up, nor would be it be fair to praise them if the rate dips.

Figure 2. Vermont's unemployment rate mirrors the national rate



Source: U.S. Bureau of Labor Statistics, Unemployment Rate and Unemployment Rate in Vermont, retrieved from FRED, Federal Reserve Bank of St. Louis; December 16, 2021.

Example: As noted in the Outcomes Report, the number of days that air quality posed a moderate or greater risk to sensitive populations is significantly impacted by natural phenomenon and emissions that originate *outside* of Vermont. Thus, it provides limited insight into how well the State is addressing air quality issues.[3]

Example: In contrast, the indicator for phosphorus loading to Lake Champlain measures phosphorous loading specifically from Vermont and largely reflects State efforts to improve water quality.

[3] The Outcomes Report also notes that this indicator is best viewed over decades, yet only seven years of data are available in the Report.

2. CLEAR GOALS AND OBJECTIVES: The population-level goals are broad and undefined, which leaves the door open for multiple interpretations of success and limits the Report's usefulness as a tool for accountability. Take, for example, the first population-level outcome: "Vermont has a prosperous economy." Do we have a "prosperous economy" if incomes rise but only among the wealthy? If paychecks aren't sufficient for many Vermonters to meet their basic needs?[4] If families can't afford health care? Defining the goal – in this case, "a prosperous economy" – is a critical first step in performance measurement with cascading implications about which indicators to use.

"How can we be prosperous if there is food insecurity?"
– State legislator

Similarly, most population-level indicators do not have a specific objective or target. In some instances, a target may not be necessary because measurable improvement is a goal in itself. When targets *are* used, it is helpful to explain what the target represents and how it was established. This is important because targets influence how we interpret findings; an arbitrary or unrealistic target could be misleading.

> **Example:** The "target" for the fall-related death rate indicator is set at 116.9 per 100,000 adults age 65 and older. Based on reading the Outcomes Report, is not clear how this target was set or what achieving this target means. (Diving in deeper into Healthy Vermonters 2020, we found that the target represents a 10% reduction from the 2009 level.)

3. SIGNIFICANT AND RELEVANT: Many indicators do not capture the most significant or relevant information. There are hundreds – if not thousands – of potential indicators that relate to the ten population-level outcomes. Selecting the most meaningful indicators is, therefore, central to the usefulness of the Outcomes Report. However, we identified a number of indicators that do not capture the most significant and/or relevant data. In some cases, adding additional detail would make the indicator more useful; in other cases, replacing the indicator would be more appropriate.

> **Example:** The first indicator listed for "Vermont has a prosperous economy" is the rate of non-public sector employment. It is not clear how – and no explanation is provided – the breakdown of public sector versus private sector employment relates to the prosperity of Vermont's economy. The graph implies that public sector employment is not as good as private sector jobs. In fact, public sector employment often creates good jobs that help provide and maintain critical infrastructure that relates directly to prosperity.[5] For instance, if the Vermont National Guard or the U.S. Customs and Border Protection branch in St. Albans were to shut down, the rate of non-public sector employment would rise, but would legislators view this as an indication that Vermont State government was delivering a "prosperous economy"? Furthermore, there is a strong positive correlation between public
>
> **"[…] Alone, it's a useless figure. It implies private sector jobs are better than public ones. It depends on what the jobs are."**
> ***– State legislator***

[4] Per 2 V.S.A. § 526, the legislature has defined "Basic needs" as the essentials needed to run a household. The Joint Fiscal Office develops the Vermont Basic Needs Budgets and Livable Wage Report biennially.

[5] Hoffer, D. The Vermont Job Gap Study, Phase 10-Part 1: Business Climate Revisited - Domestic Business Relocation and Jobs. Prepared for the Peace and Justice Center. January 2010.

sector and private sector employment, suggesting that one does not inhibit the other (see Appendix).

Rather than focusing on the breakdown of public and private sector, reporting on the quality of jobs in Vermont would tell us more about how Vermonters are doing. For example, how do wages compare to other states? Do employers offer good benefits?

Example: The rate of petitions granted for relief from domestic abuse is an indicator for "Vermont is a safe place to live." Domestic abuse poses a serious threat to the safety of Vermonters and merits the attention of policy makers. However, as an indicator, the *rate of petitions granted* does not provide the most helpful or relevant information. It is not clear whether a decrease in petitions granted is necessarily a good thing, particularly since we know that domestic abuse often goes unreported.[6] Other metrics – such as the number of calls received on the domestic violence hotline or people served by domestic violence centers – provide more straightforward insight into lives impacted by domestic abuse.[7]

"Not everyone who shows up at the shelter has sought relief in the courts, and those numbers are going *up*. The numbers don't reflect what I'm seeing at the local domestic violence shelter."
– State legislator

You be the judge: Which datapoint do you find more informative?
In 2019, there were 0.005 petitions granted for relief from domestic abuse per 1,000 residents. -OR- In 2019, the Vermont Network received 18,921 calls to the domestic violence hotline.

Example: Looking at the indicator "percentage of Vermont covered by state-of-the-art telecommunications infrastructure," you might conclude that Vermont has a robust and modern telecommunications system. According to Figure 3 below, 93% of Vermont has been covered by state-of-the-art telecommunications since 2017. Yet, we know that Vermont's telecommunications infrastructure needs significant expansion and improvement, especially regarding cellular and broadband coverage.[8] Although it is difficult to know what *is* being measured by this indicator, it clearly does not provide the most relevant information for policy discussions regarding Vermont's telecommunication needs.

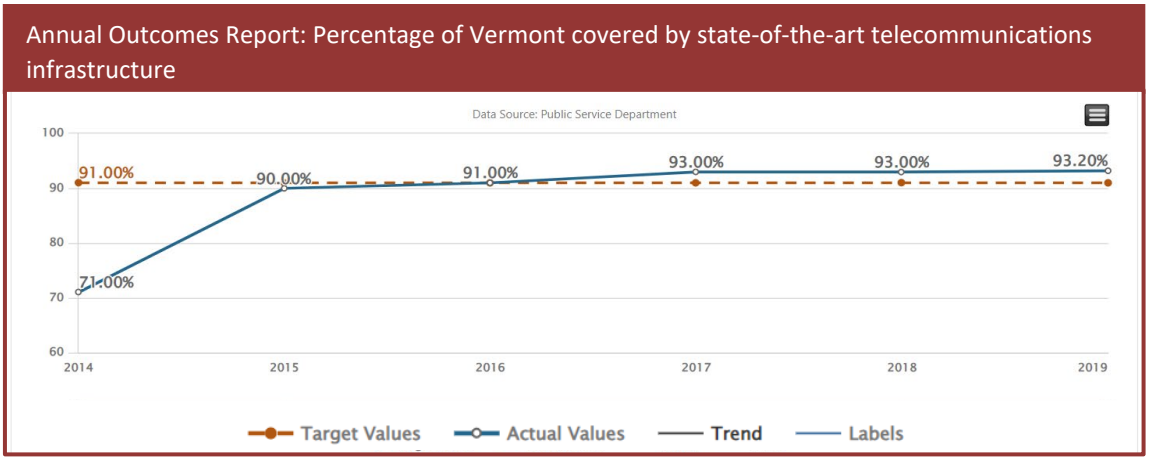
---

6 From 2006-2015, 56% of domestic violence victimizations were reported to the police. Reaves, B. Police Response to Domestic Violence, 2006-2016. Bureau of Justice Statistics. May 2017.

7 Like many indicators in the report, the prevalence of domestic abuse is influenced by factors outside of the control of State government.

8 The Vermont 10-Year Telecommunications Plan, released in the summer of 2021, notes that up to 40% of Vermont premises and up to 6,000 miles of highway may not have access to outdoor mobile voice and data, that approximately 63% of premises may not have mobile voice connectivity indoors, and that approximately 20% of premises do not receive at least 25 Mbps download and 3 Mbps upload speeds.

Figure 3. The telecommunications indicator does not reflect Vermont's cellular or broadband needs, two topics that are central to telecommunication infrastructure discussions in Vermont

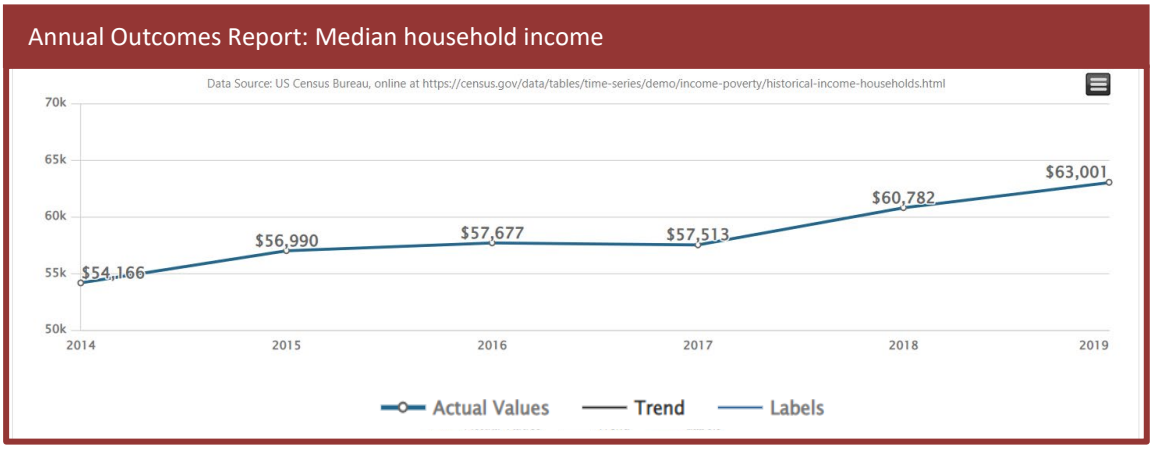


"Is this cell or broadband? I don't know. This makes us look good, but I don't think we do!"

– State legislator

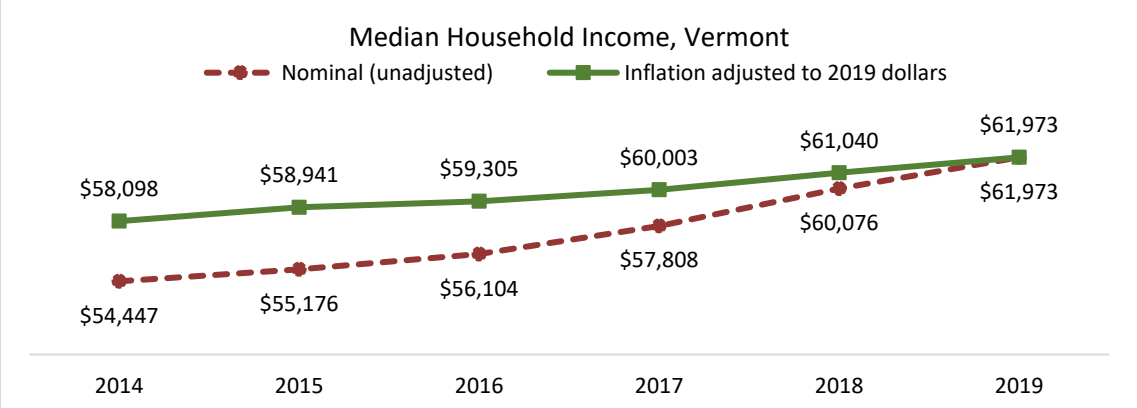
Example: Median household income is an indicator for "Vermont has a prosperous economy." Changes in income have a significant impact on Vermonters' lives; however, because the data is not adjusted for inflation, this graph does not provide the relevant information that would allow readers to assess how well Vermonters have truly fared over time. At first glance, Figure 4 shows a 16% ($8,835) increase in median household income in six years.

Figure 4. Without adjusting for inflation, it is difficult to know how Vermonters' purchasing power has changed



In Figure 5, however, we made a few adjustments: 1) we included both the unadjusted (nominal) change in income and the *inflation adjusted* change in income, and 2) we opted to use more reliable data (which is why the numbers differ from Figure 4).[9]

Figure 5. Adjusting data for inflation shows the change in Vermonters' purchasing power

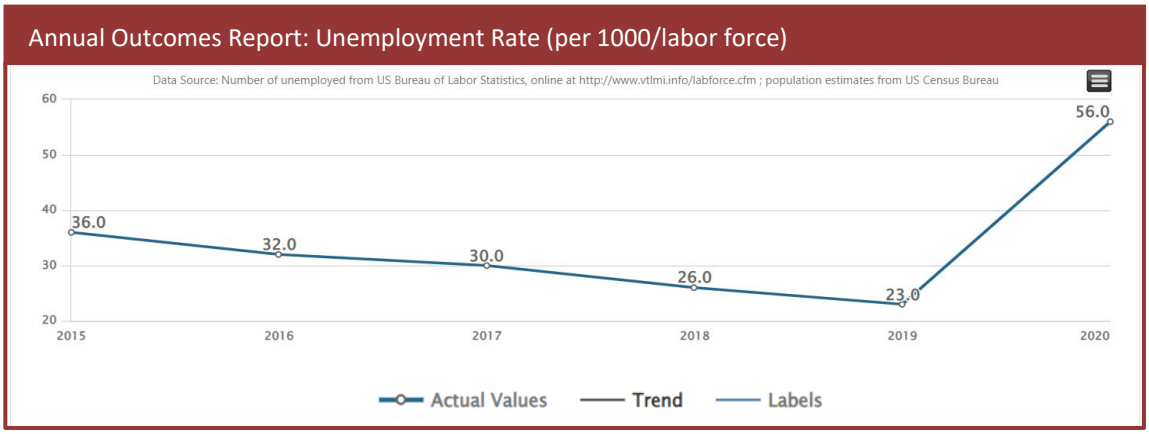


Source: U.S. Census Bureau, American Community Survey 5-Year Estimates, 2014-2019. Table S1901.
Note: Income data was adjusted using the annual Consumer Price Index for the Northeast region, All items.

Here we see that the inflation-adjusted change in median income is more modest. Instead of a 14% increase ($7,490), the inflation-adjusted median income increased by only 7% in six years ($3,875). This represents the *actual* change in Vermonters' purchasing power. Looking back ten years, inflation adjusted growth was 4% (see Appendix).

Example: The unemployment rate is another indicator for "economic prosperity." The most common measure of unemployment, called U-3, is produced by the U.S. Bureau of Labor Statistics (BLS). However, the Outcomes Report does not use the standard BLS format (e.g., Vermont's unemployment rate in 2018 was 2.6%) which makes it difficult to compare over time or against other states.[10]

Figure 6. Using a non-standard format makes it difficult for readers to interpret the data



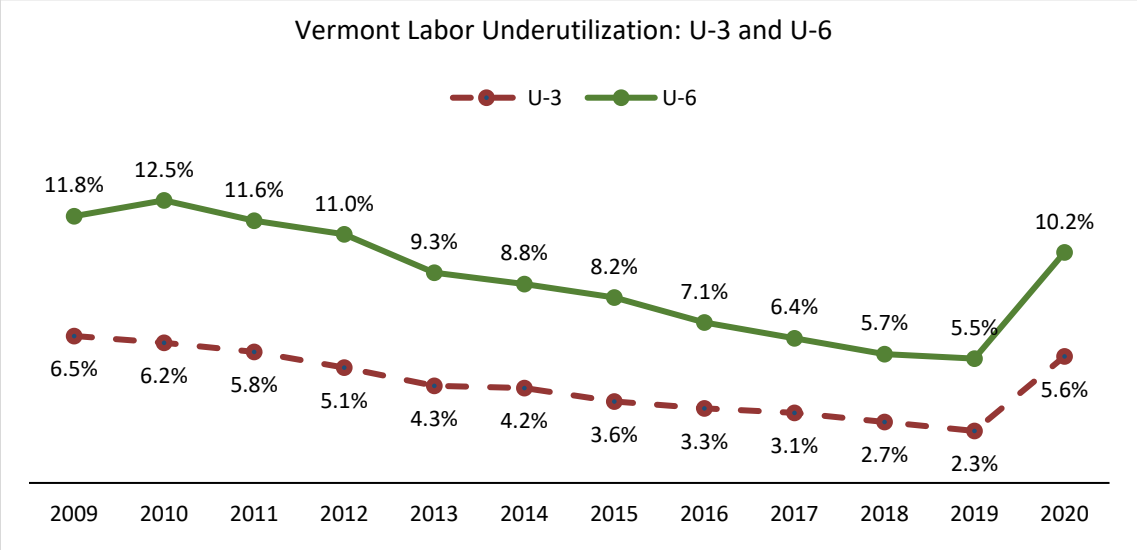
9 While the Outcomes Report uses data from the U.S. Census American Community Survey (ACS) 1-Year estimates, we opted to use the ACS 5-year estimates as they provide more reliable data. The ACS dates back to 2010; we recommend using a longer time frame where possible (see Appendix).

10 No comparative data was provided; Section 4 describes the benefits of comparative data in more detail.

Additional information could be added to this chart to provide policy makers with a more comprehensive look at the number of Vermonters whose needs are not being met by the economy. The BLS collects data on other measures of labor underutilization, such as U-6 which includes unemployed persons as well as discouraged workers, marginally attached workers, and persons who are employed part time for economic reasons.[11]

"I'm not sure from this how many people are underemployed. Those people are having a tough time making it."
– State legislator

Figure 7. Including U-6 gives a better sense of the number of Vermonters whose needs are not being met by the economy



Source: U.S. Bureau of Labor Statistics, Alternative measures of Labor Underutilization for States.

4. COMPARABLE: Most indicators lack contextual information and comparative data that would enable readers to draw meaningful conclusions. Building off the RBA framework, the Outcomes Report includes fields for departments to provide explanatory information, such as the "Story Behind the Curve," "Partners," "What Works," and "Strategy." While some agencies have filled out these sections (notably the Agency of Human Services), this information is missing for most indicators.

[11] Per BLS, unemployment (U-3) includes all jobless persons who are available to take a job and have actively sought work in the past four weeks. U-6 includes unemployed persons as well as discouraged workers, marginally attached workers, and persons who are employed part time for economic reasons. Discouraged workers are persons who are not in the labor force, want and are available for work, and had looked for a job sometime in the prior 12 months. They are not counted as unemployed because they had not searched for work in the prior 4 weeks, for the specific reason that they believed no jobs were available for them. The criteria for the marginally attached are the same as for discouraged workers, with the exception that any reason could have been cited for the lack of job search in the prior 4 weeks. Persons employed part time for economic reasons are those working less than 35 hours per week who want to work full time, are available to do so, and gave an economic reason (their hours had been cut back or they were unable to find a full-time job) for working part time.

Table 1. Many of the 56 indicators lack background information that would help the reader place the findings in context

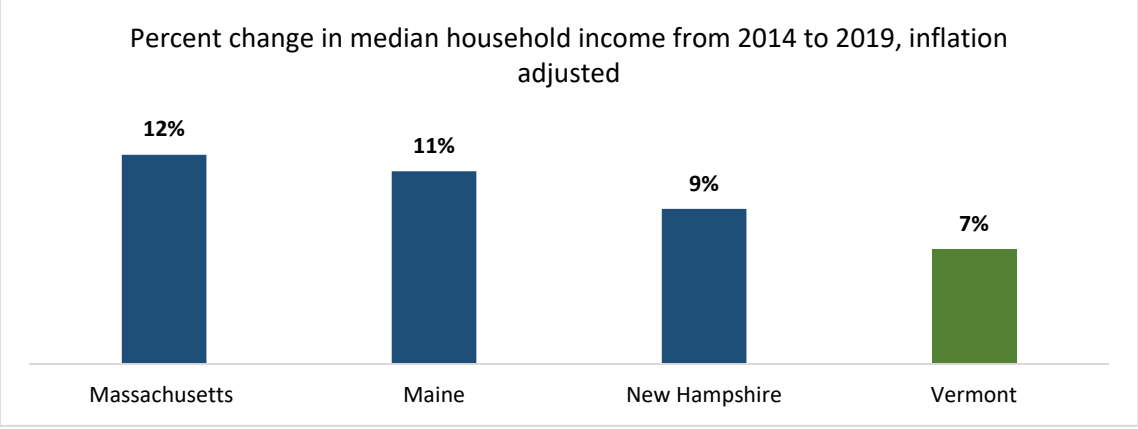
	Story Behind the Curve	Partners	What Works	Strategy
Limited detail or not filled out	50% of indicators	68% of indicators	75% of indicators	80% of indicators

Second, very few indicators include a point of comparison that would help the reader assess conditions in Vermont and, where possible, make a judgment about policies and programs that may influence outcomes. For example, providing data from neighboring states, some of whom have very different policies and strategies, would allow the reader to get a sense of the extent to which different policies make a measurable difference in outcomes or whether trends in Vermont mirror regional or national trends. In addition, providing historical data would allow for comparison over time since many population-level conditions are slow to change and are best viewed over decades. However, the graphs generally only display six or seven years.

Example: There is no background information provided for the percentage of children ready for school in all four domains of healthy development, making it difficult for readers to know what this is measuring, why it is important, what success looks like, and what the State is doing. In contrast, the Department of Health provides background information about the food insecurity indicator, including steps that they are taking and links to access more detailed information.

Example: Rather than present Vermont's median income in isolation, as the Report does, comparing it with neighboring states would help us understand how we measure up and whether the changes we see are unique to our state. Adding data from three New England states – Massachusetts, New Hampshire, and Maine – we see that from 2014 to 2019, inflation-adjusted median incomes rose in all states though the rate of growth was slowest in Vermont.

Figure 8. Regionally specific data allows the reader to compare conditions in Vermont



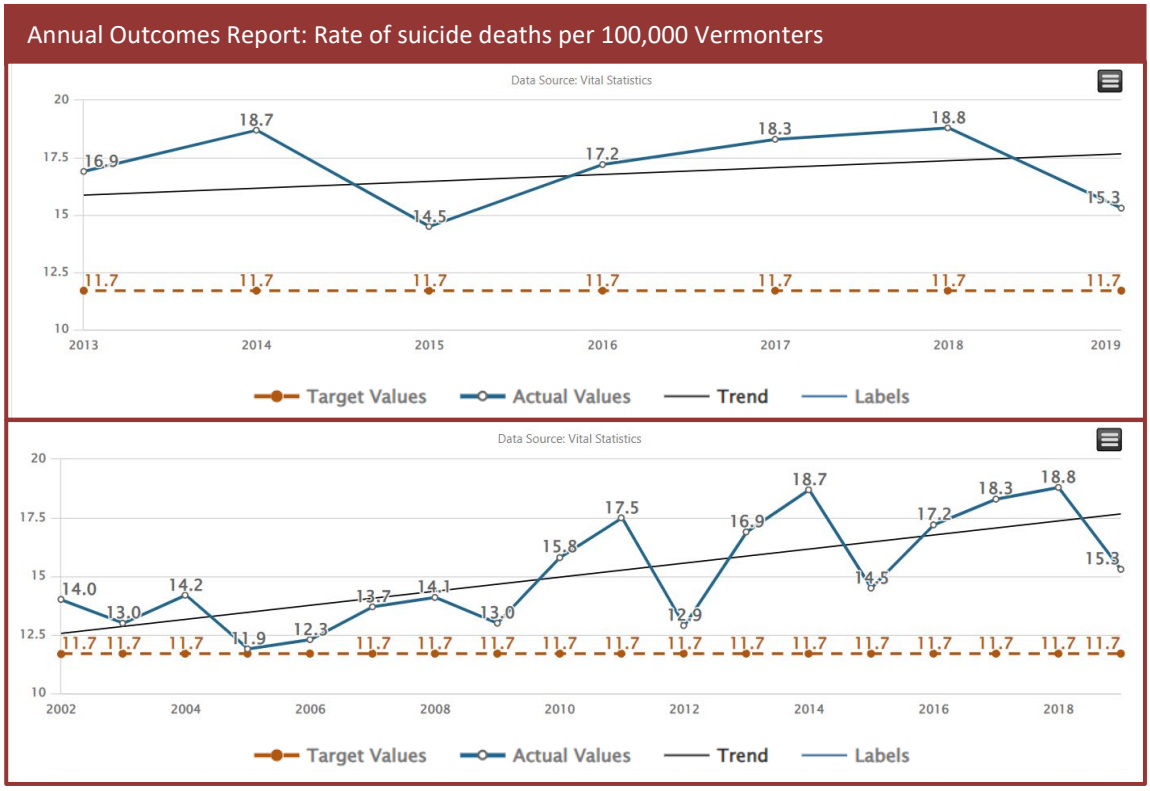
Source: U.S. Census Bureau, American Community Survey 5-Year Estimates, 2014-2019. Table S1901.
Note: Income data was adjusted using the annual Consumer Price Index for the Northeast region, All items.

Example: The "Story Behind the Curve" for the rate of petitions granted for relief from domestic abuse states: "An analysis would need to be undertaken of the Vermont, Maine and New Hampshire court systems to determine a Northern New England benchmark for this measure to

ensure an accurate comparison." At this time, however, no comparative data is provided. As we noted above (pg. 10), "rate of petitions granted" could be replaced with a more straightforward indicator – such as number of calls to the domestic violence hotline or people served by domestic violence centers – which could be compared to trends in New Hampshire and Maine.

Example: Displaying data over a longer timeframe is important when assessing population-level trends which are often slow to change. For example, the Report's default setting shows seven years of data for the rate of suicide deaths.[12] Looking at this timeframe, a reader may conclude that the rate of suicide deaths has been relatively constant. Helpfully, the Department of Mental Health provides additional years of data. When you expand to include all 18 years of data, you can see the more dramatic increase over time, as evidenced by the trend line.

Figure 9. A longer time frame is appropriate for many population-level metrics



5. DISAGGREGATED: When appropriate, disaggregating data – by income, region, gender, race, age, disability status etc. – can inform policy discussions about the disparate and unintended inequitable impact of policies and programs and/or where to target resources. The Government Accountability Committee and Chief Performance Officer, in partnership with the Agency of Human Services, have identified indicators that would benefit from disaggregated data, with a focus on race and gender.[13]

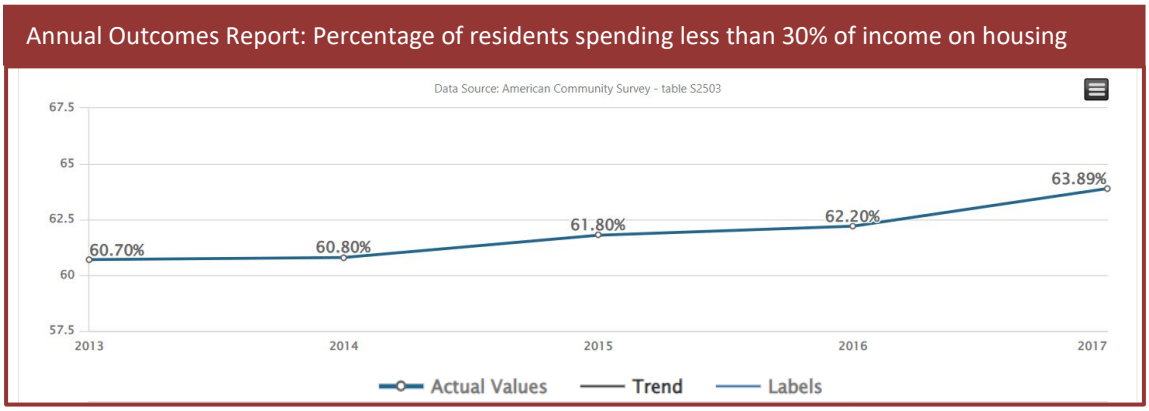
[12] Suicide is a population health measure for Vermont's All-Payer Accountable Care Organization Model.

[13] The 2021 Outcomes Report lists the indicators that can currently be disaggregated by race and ethnicity. In some cases, disaggregation of available data is not appropriate due to small sample sizes.

Data can be disaggregated in many ways to provide a more complete picture of conditions in Vermont across groups.

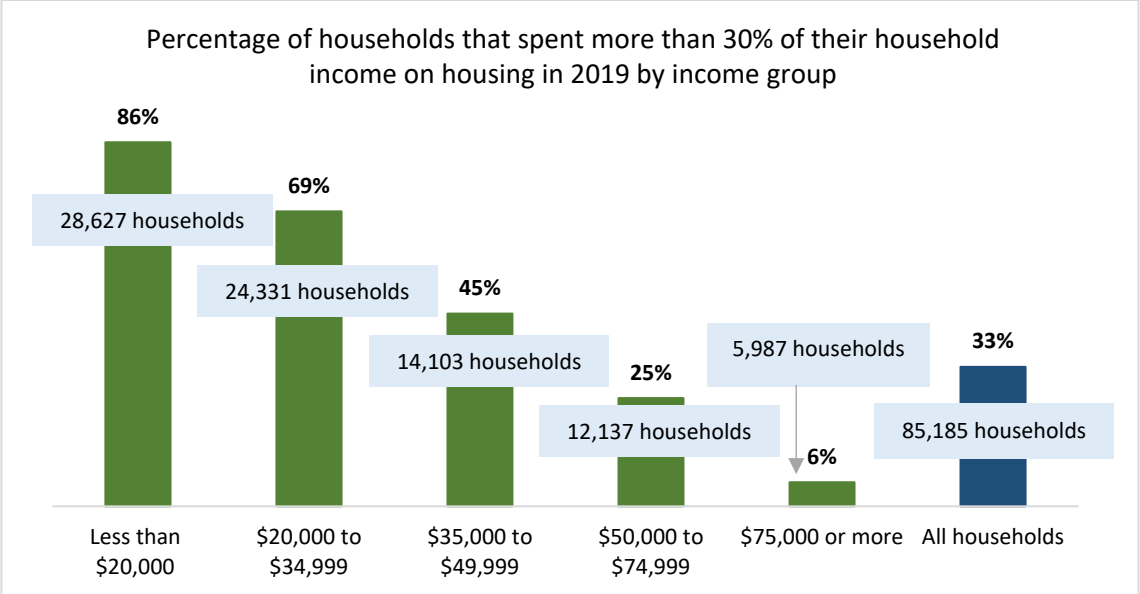
Example: Access to affordable housing is a critical issue in Vermont, particularly for low- and moderate-income Vermonters. Looking at the percentage of residents spending less than 30% of income on housing in the Outcomes Report, one might conclude that things are moving in the right direction (particularly because the limited y-axis emphasizes the small year-to-year changes).

Figure 10. Showing this data in aggregate misses the impact on low- and moderate-income Vermonters



However, when we break down this indicator by income quintile, we see – not surprisingly – that low-income Vermonters are more likely to spend more than 30% of household income on housing. This type of breakdown is critical when examining an indicator related to affordability.

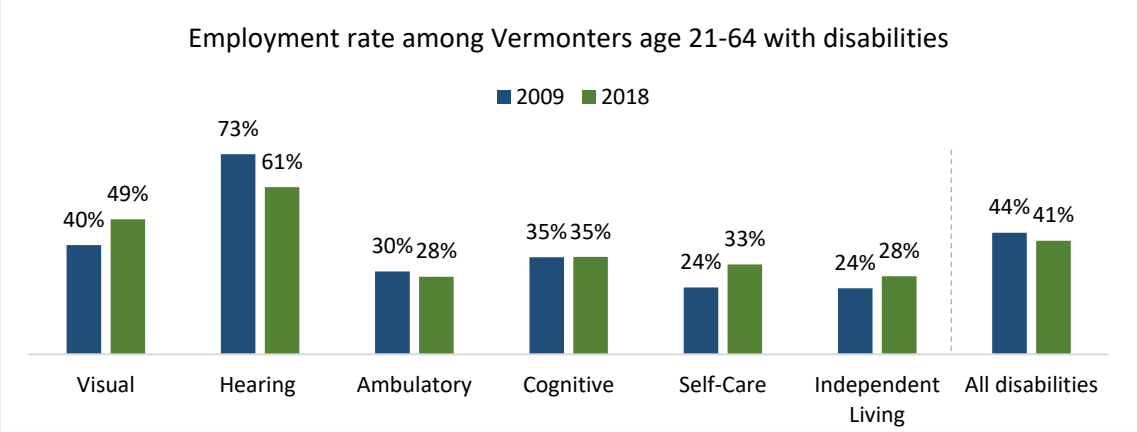
Figure 11. Low-income households are more likely to spend more than 30% of their income on housing



Source: U.S. Census Bureau, American Community Survey 5-Year Estimates, 2019. Table S2503.

Example: Disaggregating the employment rate among Vermonters age 21-64 with all disabilities by type of disability highlights the variability in outcomes by disability status.

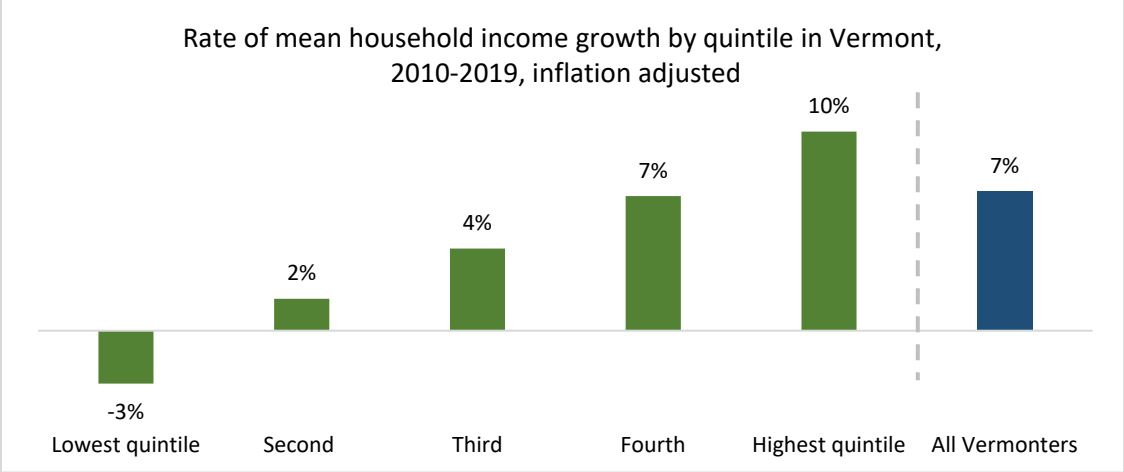
Figure 12. By disaggregating data, we can see that persons with ambulatory, self-care, and independent living have the lowest rates of employment



Source: Erickson, W., Lee, C., von Schrader, S. (2017). Disability Statistics from the American Community Survey (ACS). Ithaca, NY: Cornell University Yang-Tan Institute (YTI). Retrieved from Cornell University Disability Statistics.

Example: Disaggregating household income growth by income group tells a very different story than the aggregate data presented in the Outcomes Report (see Figure 4 on page 11). By disaggregating, we can see that from 2010 to 2019 average income *decreased* among Vermonters in the bottom income quintile but grew by 10% among Vermonters in the top quintile, telling us that growth in new income has been skewed towards higher earners.

Figure 13. Incomes have grown significantly more among Vermonters with higher incomes



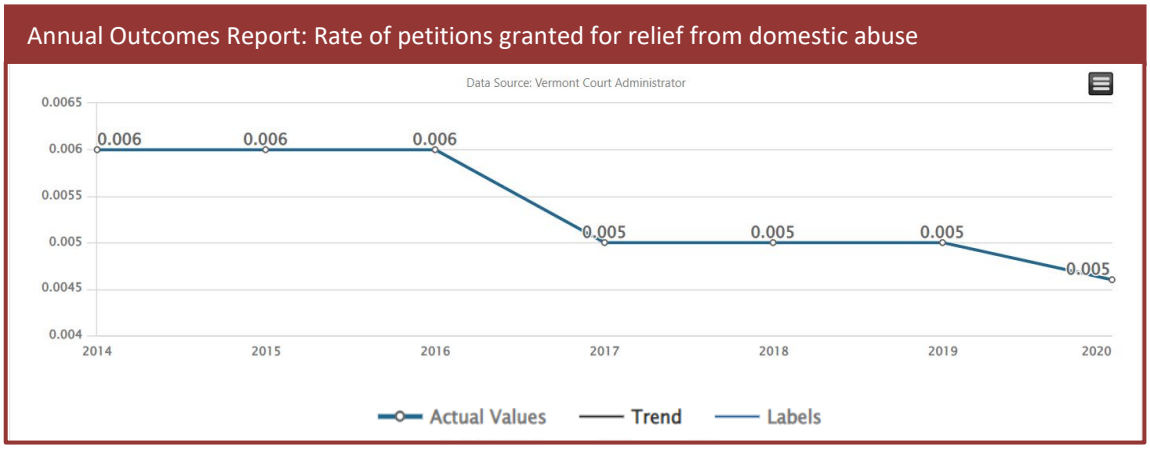
Source: U.S. Census Bureau, American Community Survey 5-Year Estimates, 2010-2019. Table B19081.
Note: Income data was adjusted for inflation using the annual Consumer Price Index for the Northeast region, All items.

"What's been the [income] growth by quartiles or quintiles? […] That's what we need to see." – *State legislator*

6. UNDERSTANDABLE: There are opportunities to improve how data is presented to make it easier for the reader to understand. The use of graphs and charts and the format of the data (e.g., raw numbers vs. percent) can have a significant impact on how readers interpret findings. Selecting the right visual format and type of data to display can make the graphs more accessible.

Example: The graph for rate of petitions granted for relief from domestic abuse illustrates two challenges that came up frequently in the Outcomes Report: scaling of the y-axis and the use of rates. First, it is often best practice, especially when working with extremely small numbers, to start the y-axis at zero to avoid exaggerating findings. As seen in the example below, limiting the y-axis from 0.004 to 0.0065 makes the drop from 0.006 to 0.005 cases in 2017 appear dramatic, when in fact, it is de minimis – a change from 6 to 5 cases per 1,000 residents.

Figure 14. Limiting the scale of the y-axis exaggerates the declines in 2017 and 2020

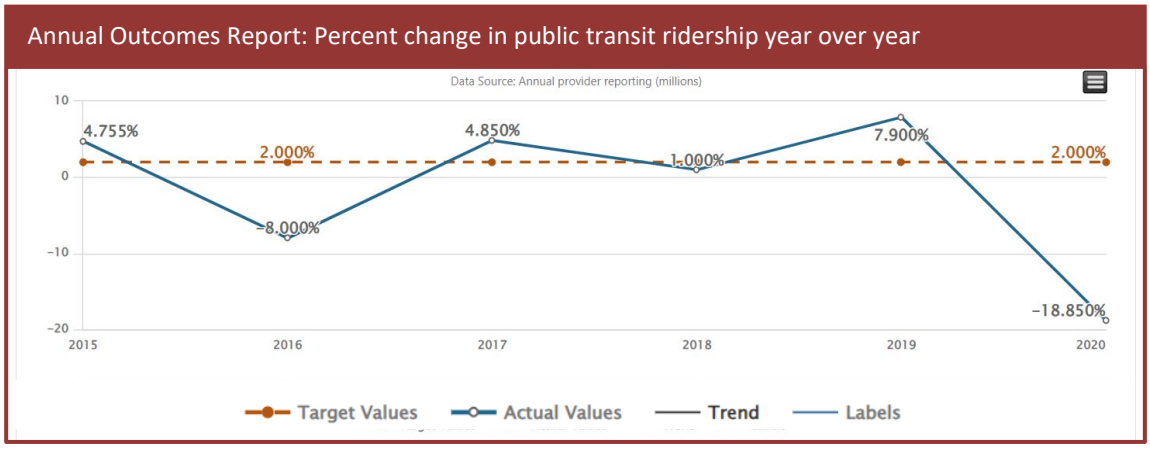


Second, while presenting data as a rate can be useful when comparing trends across settings, using a rate in this context just makes it difficult to get a sense of scale without offering any opportunities to compare the rate to other states or jurisdictions.

Example: The graph for percent change in public transit year over year is very difficult to interpret. As the only indicator on this topic, using percent change does not provide a sense of scale (e.g., depending on the total number of riders, a 2% increase could be 100 riders or 1,000 riders) and makes it very difficult to track total ridership over time. For example, how did ridership in 2018 compare to 2016? It's hard to say.

"How would you know what the number of rides that actually occurred are? This gives you weirdly fragmentary information that you can't plug into the real world."
– State legislator

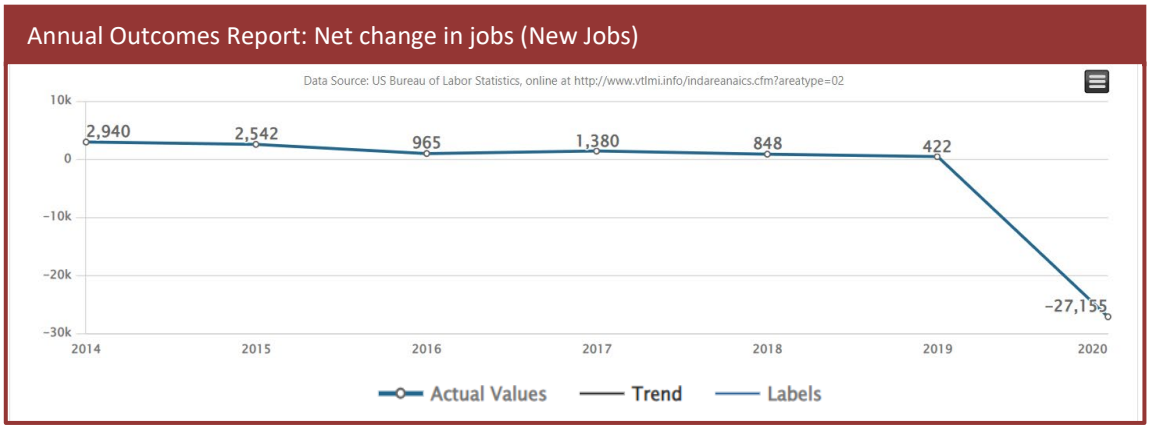
Figure 15. While an important topic, the presentation of the data is confusing and misses critical information



It is also counterintuitive that a 2% increase is depicted as a straight line. Showing the number of riders with an annual target would be much easier to interpret. Even better, the Agency of Transportation could provide a breakdown by type of public transportation and region.

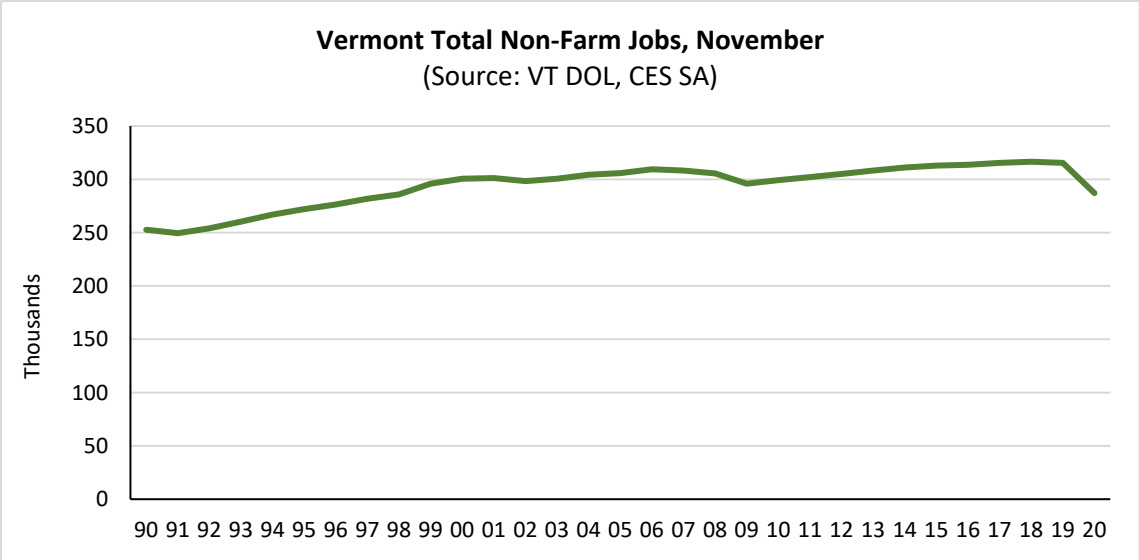
Example: Looking at the net change in jobs indicator, it is difficult to get a sense of the magnitude of these changes.[14] In contrast, displaying the data cumulatively over a longer time period allows the reader to place the findings in context. For example, Figure 17 allows you to compare the decline in jobs following the 2008 recession with the impact of the COVID-19 pandemic.

Figure 16. It is difficult to interpret the scale of the changes in this graph



[14] It is also worth noting that the "net change in jobs" indicator does not differentiate between public and private employment (see page 9). In a given year, public jobs gained could exceed private jobs lost. In that case, you would see a positive change in jobs, but a decline in the rate of non-public sector employment. This example highlights the need to consider the relationship between the various indicators.

Figure 17. Displaying the change in jobs cumulatively places the 2020 drop in context



Source: U.S. Bureau of Labor Statistics, Current Employment Statistics, State and Metro Area.

7. WELL-DEFINED: Some indicators have no definitions of the variables being measured, without which the reader is left guessing.

Example: It is not clear what is being measured by the percentage of Vermont covered by state-of-the-art telecommunications infrastructure. What does state-of-the art-mean? What type of telecommunication infrastructure? Land line? Cell phone? Broadband? Are they measuring coverage by population? Residence? Geography?

"What does 'state-of-the-art mean? Are they saying your cell connection meets your needs? Fiber?"
– State legislator

Example: In contrast, the indicator for the percentage of adults age 18-24 binge drinking in the last 30 days provides details about the group being measured, the timeframe, and the activity (binge drinking is defined in the "Notes on Methodology" section).

8. VALID: Many indicators lack specific data sources or methodology information, making it difficult for the reader to assess the validity of the information. In some instances, links to sources are provided but do not work or are not clickable. Furthermore, in multiple places, we found mislabeled or inconsistent data.

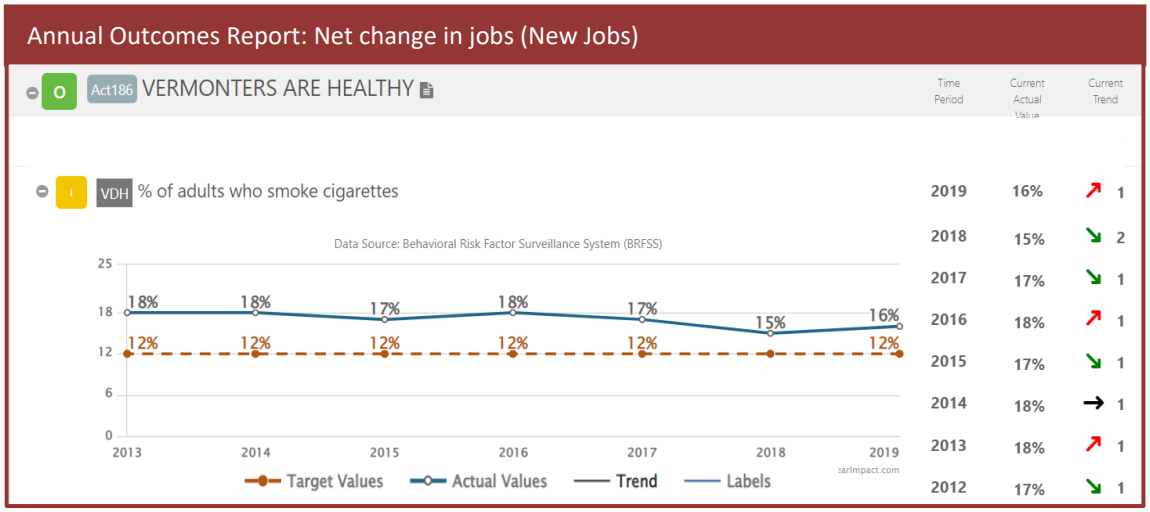
Example: No source is provided for the percentage of children below the basic level of fourth grade reading achievement under state standards or the Gross State Product per capita.

Example: The rate of violent crime per 1,000 crimes is an indicator for "Vermont is safe." The data source is listed as "FBI and VCIC," but no links are provided. We were able to find the source data on the Federal Bureau of Investigation's Crime Data Explorer page; however, upon reviewing their data, we found that the graph in the Annual Outcomes Report is mislabeled. The FBI data measures the rate of violent crime by *population*, not as a rate per 1,000 *crimes*.[15]

"Per '1000 crimes'? It doesn't give me any information I need to know what 'violent crime' means and what else is happening with other crimes."
– State legislator

Example: In addition to the graphs for each indicator, the Outcomes Report includes a summary table that displays color coded "Current Trend" arrows that show the directional change since the previous period. The number next to the arrow indicates how many periods this trend has been continuing. For example, in the graph below, the rising arrow next to 2019 indicates that smoking increased in 2019 (colored red because this is not the desired direction) and that this trend has occurred for one period.

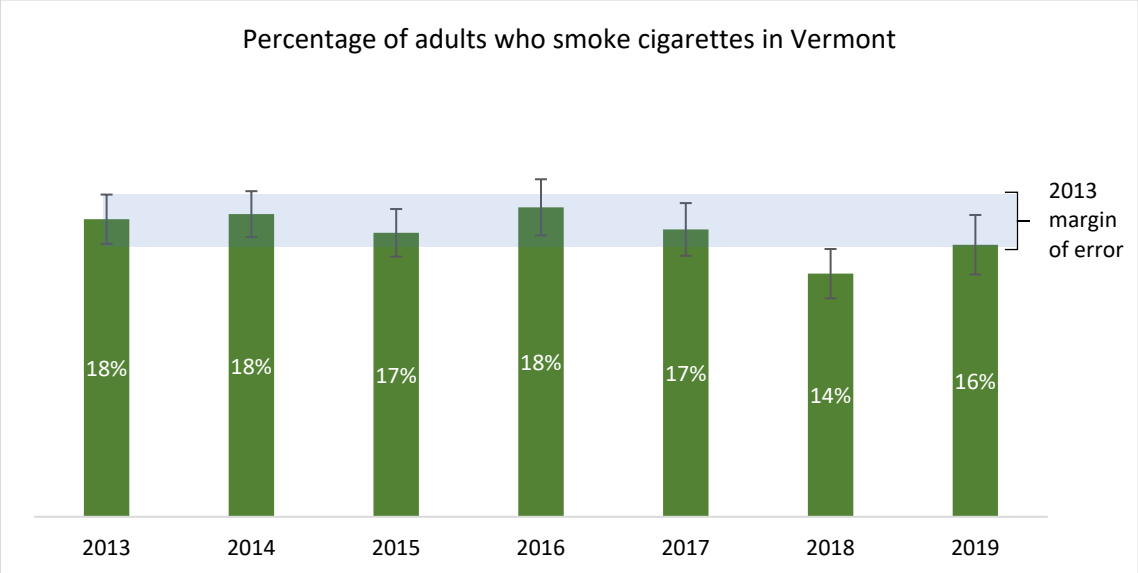
Figure 18. Presenting small changes within the margin of error as "Trends" may be misleading



While the arrows suggest that there have been meaningful changes in the trends over time, diving deeper into the data, we found that the small year-to-year changes in smoking mostly lie within the margin of error for other years. In other words, we can't say with confidence that anything has changed in most years.

[15] The FBI data on violent crimes is presented as violent crime offenses per 100,000 people. In the Annual Outcomes Report, the rate of violent crime is presented per 1,000 people (mislabeled as per 1,000 crimes). Keeping the data in the FBI format (per 100,000 people) would allow for Vermont's data to be compared with other states more easily.

Figure 19. With the exception of 2018, every data point is within the margin of error



Source: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Division of Population Health. BRFSS Prevalence & Trends Data. 2015. Accessed Jan 10, 2022.

9. TIMELY: The indicator data is generally current, though some descriptions and targets need updating. In some instances, the most recent available data date back several years due to data collection or analysis lags. This is an understandable limitation that cannot always be avoided. Other data, targets, and accompanying text need to be updated to the most current year.

> **Example:** Although the data for the percentage of Vermont adults with any mental health conditions receiving treatment has been updated more recently, the "Story Behind the Curve" has not been updated since January 2017.

> **Example:** In addition to regularly refreshing the data, the indicators and targets should also be updated. For example, the definition of "state-of-the-art" telecommunications has evolved over time as technologies change, but the indicator and the target have not been updated to reflect this.

Conclusion

The Annual Outcomes Report is intended to provide legislators and the public the data that they need to judge whether State government is doing a good job. Now in its eighth year, our analysis uncovered several challenges that limit both the quality and usefulness of this Report. These findings were underscored during our conversations with legislators (most of whom had not seen the report in years – if at all). The magnitude of the issues that surfaced in our analysis ranged from fundamental to technical. On the one hand, the question of

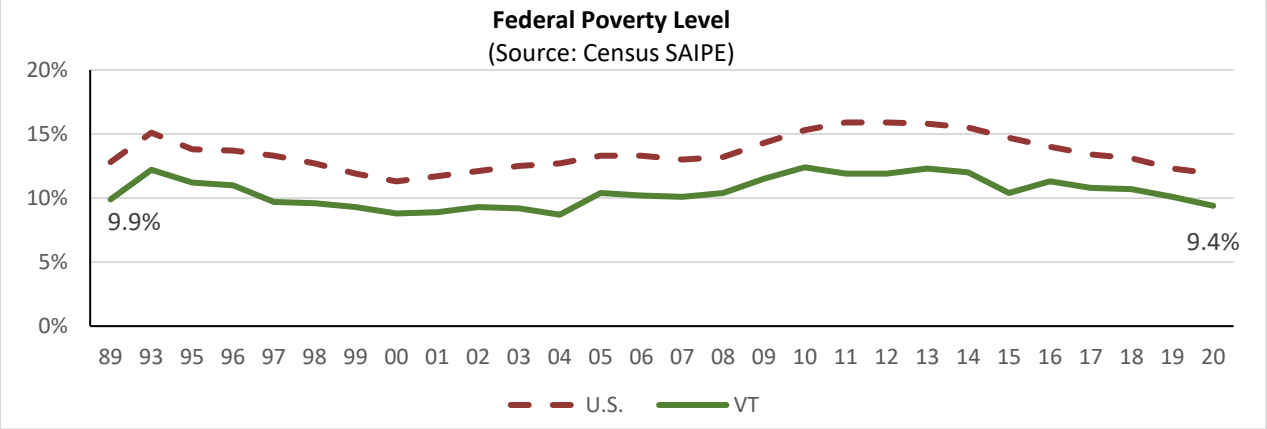
"[The Government Accountability Committee should regularly] go back and scrub [the Outcomes Report]. Ask whether we are getting what we want from these indicators."

-State legislator

"influence" – whether the State can make a meaningful impact on population-level outcomes– raises deeper questions about the usefulness and appropriateness of the RBA population accountability framework, as well as the Annual Outcomes Report itself. (It's worth noting that the use of RBA-style population-level goals and indicators is not unique to the Outcomes Report; we found many of the same issues in the Governor's Strategic Plan.) On the other hand, many of the data presentation issues could be resolved more easily given the right resources and skillset. We hope the Chief Performance Officer and the Government Accountability Committee make use of our analysis and the new framework we offer as they consider the future of the Report.

Despite these challenges, the Outcomes Report does contain important information about conditions in Vermont. Take, for example, the Federal Poverty Level (FPL), one of the indicators for "Vermont has a prosperous economy." Looking across the last thirty years, the FPL has remained largely unchanged in spite of billions of federal and state spending. This signals that the combined force of current public policies and funding is not meaningfully impacting what is clearly a key statewide indicator – the percent of Vermonters living in poverty. This should be a cause for reflection. It highlights the need to put individual indicators in perspective.

Figure 20. The percentage of Vermonters living in poverty has remained relatively flat[16]

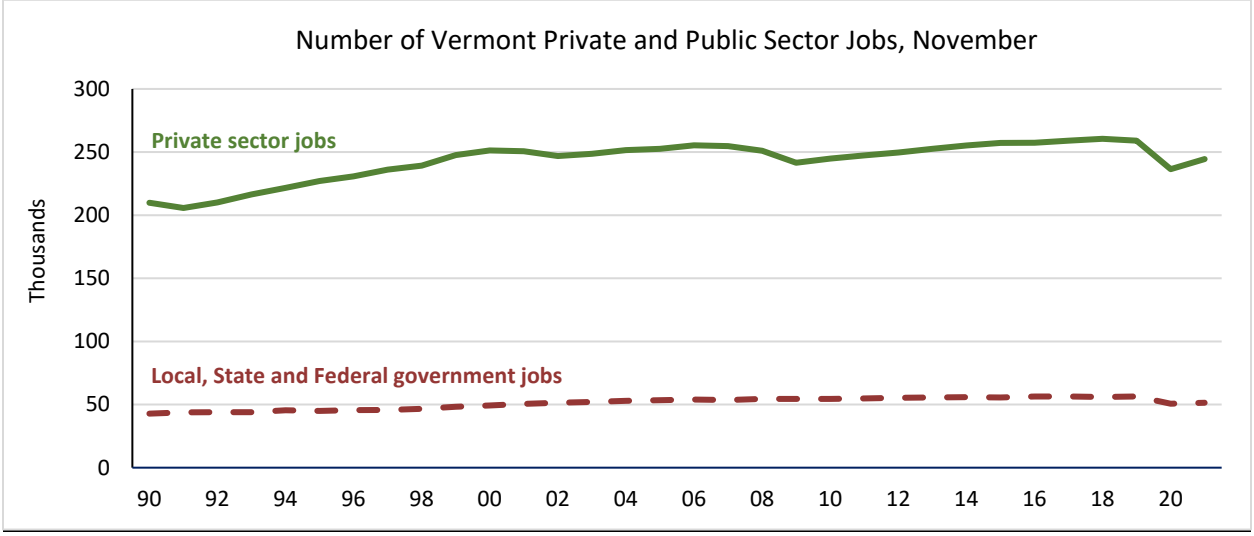


On a final note, it is important to remember that producing high quality performance data is one half of the equation. A performance report can be great and still collect dust. High quality performance measurement must be paired with a commitment from legislators and administrators to use the data to drive decisions that impact real people.

[16] We opted to use the percentage of Vermonters at 100% FPL – rather than 185% FPL which is used in the Outcomes Report – because data was readily available for a longer time period.

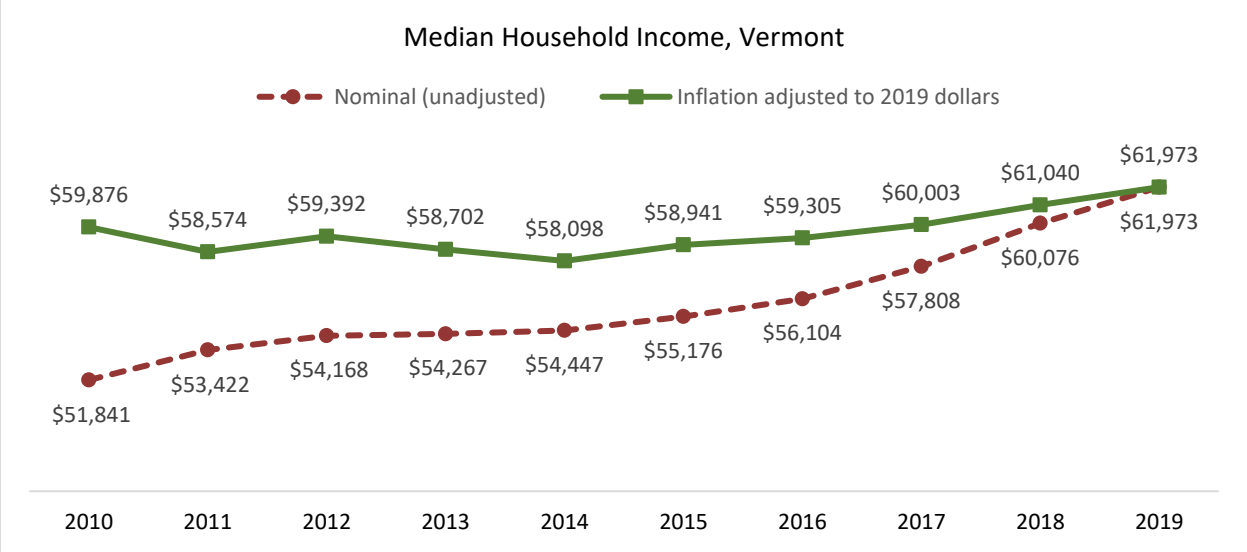
Appendix

PUBLIC AND PRIVATE SECTOR JOBS: Public sector jobs have constituted 16% to 18% of Vermont's total jobs for thirty years. Over the years, there has been a very strong positive correlation between the two.



Sources: Vermont Department of Labor; U.S. Bureau of Labor Statistics, Current Employment Statistics, State and Metro Area.

MEDIAN INCOME: Using inflation-adjusted income over a longer time period gives better insight into the change in Vermonters' purchasing power over time. The unadjusted growth from 2010 to 2019 was 20%; however, inflation adjusted incomes grew by just 4% during that time.



Source: U.S. Census Bureau, American Community Survey 5-Year Estimates, 2010-2019. Table S1901.
Note: Income data was adjusted using the annual Consumer Price Index for the Northeast region, All items.

Methods

FEEDBACK FROM LEGISLATORS: To gather feedback on the extent to which legislators feel well-served by the Annual Outcomes Report, we reviewed a sample of indicators in one-on-one sessions with a bipartisan group of legislators from both the House and the Senate. We intentionally did not include any legislators who currently sit on the Government Accountability Committee.

TABLE 1. (pg. 14): Table 1 provides an overview of how many indicators include information about the "Story Behind the Curve," "Partners," "What Works," and "Strategy" sections. While some indicators had additional sections, these four sections were the most commonly used throughout the Outcomes Report. We reviewed all 57 indicators but excluded the Genuine Progress Indicator as the notes indicate that it is no longer being updated. In some instances, relevant information was listed in a different section. For example, partners may be included in the "What Works" section. For the purposes of our analysis, we recategorized relevant information to give credit even if the information was not labeled correctly.