USING LINKED ADMINISTRATIVE DATA TO INFORM POLICY AND PRACTICE: THE WISCONSIN EXPERIENCE

THE WISCONSIN ADMINISTRATIVE DATA CORE
The Wisconsin Administrative Data Core is a multi-system data resource designed to support the integrated analysis of multiple program participation and outcomes of individuals and families. Developed and maintained by the Institute for Research on Poverty (IRP) at the University of Wisconsin-Madison in collaboration with its Wisconsin state agency partners, the Data Core links a large array of cleaned and harmonized administrative data covering a long period of program history (see Table 1).

The system relies on a file known as the Multi-Sample Person File (MSPF), which contains one observation per individual, with no individual appearing twice. The MSPF can be linked with program participation data files, as well as complementary files that allow researchers to group individuals by case and/or by various definitions of family, and follow them over time. The 2017 MSPF includes information for just over 7 million individuals. IRP rebuilds the data source annually; thus, allowing for continual improvements in IRP’s ability to match individuals across systems, drawing on any new data collected. The system, which evolved from a series of large-scale evaluation projects conducted by IRP for the State of Wisconsin starting in the 1980s, is designed to support research that has the potential to inform the evaluation and administration of public policies and practices, and allows for cross-program comparisons and analysis that would not be possible using only a single agency’s data.

“YOURS, MINE, AND (INCREASINGLY) OURS”
How did this rich data source come together, and how might other states and localities replicate it? Certainly, one must navigate the unique technical challenges of harmonizing data from disparate sources1, but perhaps even more important is the development and navigation of the partnerships between data providing agencies and research organizations. In

Table 1: Programs/Systems in the Wisconsin Administrative Data Core*

<table>
<thead>
<tr>
<th>DEPARTMENT</th>
<th>DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Instruction (DPI)</td>
<td>Student data from WI public (including charter) schools (2005–6 – current)</td>
</tr>
<tr>
<td>Corrections (DOC)</td>
<td>Incarceration in state prison facilities (1990 – current)</td>
</tr>
<tr>
<td>Milwaukee County Sheriff</td>
<td>Incarceration in Milwaukee County jails (1990 – current)</td>
</tr>
<tr>
<td>Court Records</td>
<td>21-county sample of divorce and paternity cases (1980–2009; additional cohorts being added)</td>
</tr>
</tbody>
</table>

*Data from other sources, such as Vital Statistics, have been matched to the Data Core on a per project basis.
Wisconsin, this has meant fostering a "logic of collaboration" (See Figure 1) that supports both policy development and academic research, and recognizes that such partnerships require infrastructure and resources to support sustained engagement. For example, to create the data system, the State of Wisconsin agencies provide data; IRP provides specialized programming staff, technical support staff, and specialized hardware and software; and funding is provided by UW, grants and contracts from federal sources (often in response to joint UW/Agency proposals), state grants and contracts, and foundations.

With this engagement, partnerships can develop trust, and a shared understanding of useful and interesting questions, appropriate methods, and satisfactory answers. Although all projects using this wealth of data must have core policy or practice issues and questions as a basis for the research, and must be approved by State agencies, UW has worked in partnership with agencies based on a philosophy of “Yours, Mine, and Ours.” Specifically, partners may identify questions and projects that are of primary interest to the agency, but require the expertise of IRP researchers to answer (yours); projects that may result in generalizable learning, but are of primary interest to IRP researchers (mine); and, importantly, and increasingly, a set of questions and projects that are of high interest to both researchers and agencies, and relevant to each other’s missions (ours).

The Data Core, which provides a unique resource for agencies that cannot otherwise link and analyze across systems, sustains state support of data access and funding; and given its uniqueness as a resource for research that cannot otherwise be completed, sustains commitment and interest of academic researchers. Importantly, state agencies are not permitted to provide data access for research not relevant to the agency’s mission, so researchers need to accept these limitations, and understand and explain the utility of their research. In addition, research results must be made public; IRP faculty and staff are not permitted to submit research for clearance by funders; and they are not allowed to return matched data to agencies for legal reasons. Therefore, agencies must value and accept independence.

To identify questions of interest and support mutually beneficial research, IRP staff and faculty affiliates engage regularly with agency leadership and staff; for example, by:

- Participating in regular briefings and opportunities to check in;
- Developing sustained research agreements for programs (e.g., child support, child welfare);
- Fielding contracts for specific programs and projects (e.g., educational disparities for children in out-of-home care, Child Support Noncustodial Parent Employment Demonstration evaluation);
- Participating in quarterly “Learning Exchanges” with agency leadership and staff; and
- Providing ad hoc technical assistance to each other (some state funded, some not).

LESSONS LEARNED

The work described above has led to a number of important “lessons learned,” which may be applicable to other states and localities looking to build linked administrative data resources—and, crucially, research partnerships with state and local agencies:

1. Administrative data are generally not designed for research.
   • Most administrative records are not designed with a strict requirement to maintain a single record per individual, across time, and, especially, across programs.
   • Each program privileges the data most critical to its mission; thus, there is no commonly agreed-upon set of high-quality identifying variables, and agencies with distinct missions face challenges in determining which data source takes precedence or how conflicts should be resolved.
• Data sets change over time and historical records of these changes need to be maintained.

• Access to the data is predicated on research being of interest to the data custodian, and must be in compliance with established statutes and regulations.

2. **Integrated data system construction requires answering questions that affect the cost of maintaining the system and how the data can be used, for example:**
   • Whether to use samples or the universe of individuals?
   • How to structure the data: by individual, case, family unit, or child/adult status?
   • What are useful identifying variables?
   • How to pre-process or “clean” administrative data?
   • How to address considerations of legally accurate data, “fuzzy” matching, over-matching, under-matching, and cross-matching?
   • Which matching technique to use (deterministic, probabilistic, or a mix of the two)?
   • How and when to update the data system?

3. **Collaboration with data custodians and maintenance of the partnership are essential.**
   • Involvement by an entity that does not “own” any of the programs or the associated data can facilitate probabilistic matching.
   • Infrastructure requires sustained commitment by all parties and significant funding; big “fixed” costs are a challenge to fund and manage.

**THE PAYOFF: EVIDENCE-INFORMED POLICYMAKING**

Although there are significant technical, financial, and relationship challenges involved, investing in researcher-practitioner partnerships and developing linked data systems can lead to a culture of evidence-informed policymaking that is both beneficial to society and rewarding for researchers. Recent projects resulting from the Wisconsin partnership include:

• **Educational outcomes for children in foster care**
  — A project resulting from an agency request for technical assistance that resulted in the first opportunity to link child welfare and educational data in Wisconsin, and suggested that permanency alone is insufficient to promote foster youths’ educational and economic attainment.

• **Understanding the interaction between child welfare and child support collection**
  — A project resulting from a DCF staff question that illuminated unintended negative consequences of policy requiring parents to pay child support to offset the costs of their children’s stay in foster care.

• **National Child Support Noncustodial Parent Employment Demonstration (CSPED)**
  — A randomized controlled trial combining administrative data from eight states to test the effectiveness of child support-led employment programs for noncustodial parents.

Each of these projects has resulted in both policy innovation and academic research products; thus, truly realizing the underlying “yours, mine, and ours” philosophy at the heart of these efforts.

---


