Evaluating MEP Evaluation

Daniel Luria, PhD
VP Research
Michigan Manufacturing Technology Center (Michigan's MEP affiliate)

Why Me?

- PhD economist; major field = Economic Survey Methods (Univ. of Mich.'s Institute for Social Research)
- Longstanding interest in Center and MEP system performance; co-wrote 1991 proposal to win Michigan's cooperative agreement, which included a client-control evaluation approach
- Built the MMTC's performance Benchmarking Service (PBS) in order to have a control group for such evaluations
- Part of MEP's Evaluation Working Group, 1993-97
- Frequent collaborator with Eric Oldsman (Nexus Associates) on evaluations of NY, PA, and non-US manufacturing extension programs

What MEP Evaluations Have Found

NIST MEP's Synovate/Turner/Fors Marsh surveys of most clients 6-12 months after end of delivery of MEP services, asking for impacts resulting from services:

- New or retained sales
- New or retained sales
- Cost savings or avoidance

… but also zero or near-zero medians on every metric

Rigorous client-control studies:

- 1987-92 (Jarmin & Jensen, Census CES, 1997): no sales effect, 5% productivity (VA/employee) effect (*Note: PA IRCs dominated client roster as of 1992*.)
- 1992-97, PA IRCs only (Oldsman, Nexus Assocs, 2003): the same
- 1997-2002, SRI/GaTech (2010): no sales or productivity effect

MEP/Synovate/Turner/Fors Marsh Survey

- Stable, consistent instrumentation and questions for more than 10 years. Centers are used to it.
- Inevitable problems of client ascription of impacts vis-à-vis the counterfactual of not having received MEP services.
- Compounded by questions that, to answer meaningfully, require dozens of calculations (e.g., labor, material, overhead, & inventory costs).
- Large sum-of-impacts results, *which we know are driven by outliers because …*
  - … On all metrics, and at almost every center, zero medians
  - Zero median result consistent with SRI/GaTech study finding of no significant sales or productivity impact
  - Zero median finding does not prove that MEP "doesn't work," e.g., in clinical trials, a med with 40% effectiveness versus a 15%-effective placebo is deemed to be of demonstrated clinical value.

Survey Overestimates Sum-of-Impacts … but *Under*estimates Median Impact

- Survey design guarantees that firms with the same experience may report huge, small, or zero impacts. *Example: Center helped client achieve compliance to ISO 9000 standard required by customers accounting for 80% of its sales. Client A credits services with retaining 80% of its sales. Client B reasons that it would have achieved compliance somehow without MEP, and reports no impact. Client C views the impact as the small cost savings associated with having used MEP versus a private consultancy . Client D reports what it paid the center as an investment impact.* ***Client A generates an outlier; client B depresses the median.***
- True role of outliers hard to assess, because survey looks only at changes, with no reference to base levels. *Thus, a $1-million client reporting a $2 million impact is accepted, while a $100-million client reporting a $25 million impact has to be investigated.*
- Lack of a systematic guide to how to think about impacts (e.g., embedding instructions and worksheets in the survey):
  - Opens the door to centers coaching clients on how to respond, and
  - Almost certainly leads many clients not to quantify impacts.
- Many centers defend not fixing the survey because they need outliers to look good.

Evaluating Growth & Innovation Projects

- Nothing special about these kinds of projects. Credited new sales is the right metric, & recent survey revision gets at the drivers of those new sales.
- But there are two elephants in the room:
  1. It is highly unlikely that NIST MEP or centers can get most clients to ascribe new sales to MEP services. *Econometric studies (including MMTC's 2005 study for NIST MEP) show that we can explain about 20% of variance in productivity and profitability, but less than 5% of variance in sales growth or decline.* Nick Bloom's UK work suggests that the main drivers are unmeasured characteristics of managers.
  2. Even more serious, *all sales impacts must be presumed to be zero-sum or very nearly so for US manufacturing*. This presumption must also extend to clients' credited export sales, which must be presumed mainly to displace sales by other exporters.

Note that this presumption does *not* apply to productivity growth, where one firm's increase does not imply other firms' decrease.

What *Should* We Want to Know to Evaluate MEP?

1. **Do MEP center services make US manufacturing larger?**

   Correct metric = change in net **value-added**

   - To measure it requires:
     - Sales post- vs. pre-services
     - *Less*: Purchased inputs pre- and post
       - Percent of pre- and post- purchases from US sources

   Multiplied by import/domestic production ratio for client's 6-digit NAICS code

2. **Do MEP center services increase clients' productivity?**

   Correct metric = change in **value-added per FTE**

   - To measure it requires:
     - Same data elements as for value-added, above, plus …
     - … FTEs, pre- and post-services

Conclusions

- Current system has been logical and consistent:
  - Works passably well to defend MEP by getting large-seeming sum-of-impacts
  - Almost certainly motivates centers to strive for large impacts
- Basing claims of MEP impact on anything but change in VA and VA/FTE invites reasonable presumption of near-zero net impact.
- Current system:
  - Does not address what should matter most: *whether MEP makes US manufacturing larger or clients more productive*
  - Provides scant help to NIST MEP or its affiliated centers about *what to do more or less of to increase impact*: measurement error results in almost no association between the particular service and the predicted impacts.
- Proposed CORE approach does not remediate deficits in economic impact survey, and adds weight to measures of activity that as yet have no demonstrated relationship to economic outcomes.