

MSP  
#3**IRS Performance Measures Provide Incentives that May Undermine the IRS Mission****RESPONSIBLE OFFICIALS**

Steven T. Miller, Deputy Commissioner, Services and Enforcement  
 Beth Tucker, Deputy Commissioner, Operations Support  
 Rosemary D. Marcuss, Director, Office of Research, Analysis and Statistics

**DEFINITION OF PROBLEM**

The IRS's organizational performance measures focus on the number of activities completed during a short-term measurement period (*e.g.*, cycle time). For example, over two-thirds of the collection measures that appear in the Small Business/Self-Employed (SB/SE) division's quarterly performance reports improve if SB/SE reduces cycle time, as do nearly half of the measures reported in the IRS-wide performance appraisal. Thus, IRS leaders face incentives to adopt policies that reduce program-level processing times, even if those policies produce premature closures that do not resolve taxpayers' problems or promote the IRS's underlying mission to improve voluntary compliance.<sup>1</sup>

In addition, the IRS estimates the direct revenue it receives from enforcement, but not from taxpayer services. This practice makes it easier for Congress to fund initiatives that *generate* direct enforcement revenue than those that *prevent* delinquencies. Because the likelihood of receiving additional funding increases along with a program's direct enforcement revenue, IRS leaders face incentives to adopt policies that maximize such revenue, even if other policies would bring in more revenue by preventing delinquencies.

The incentive for IRS leaders to adopt policies that maximize processing speed and direct enforcement revenue may undermine the IRS's service-oriented mission at a time when Congress is asking the IRS to administer complex social programs, such as health care, where service is critical. Even in connection with the IRS's last reorganization – when it had responsibility for fewer social programs – legislators mandated balanced measures because they determined an excessive focus on short-term productivity and efficiency was inconsistent with the IRS's mission.<sup>2</sup> While it can be difficult to measure short-term results without creating incentives that undermine long-term goals, the IRS could do more to create incentives for IRS leaders to adopt proactive policies.

<sup>1</sup> The IRS mission is to “provide America’s taxpayers top quality service by helping them understand and meet their tax responsibilities and by applying the tax law with integrity and fairness to all.” Internal Revenue Manual (IRM) 1.1.1.1 (Mar. 1, 2006).

<sup>2</sup> See, *e.g.*, Report of the National Commission on Restructuring the IRS, *A Vision for a New IRS* 21-22 (June 25, 1997) (“the [IRS’s] internal measurement systems ... do not measure long-term quality performance accurately. Consequently, employees put an emphasis on short-term performance and meeting goals of efficiency ..., rather than on a balanced focus on efficiency, quality, and taxpayer service. Many employees and outside observers believe the result is that the performance measures do not align with the ultimate objectives and mission of the IRS.”); IRS Restructuring and Reform Act of 1997, H.R. Rept. No. 105-364, at 47 (1997) (“The Committee encourages the IRS to redesign its performance measures to more appropriately align employee behavior with organizational goals... While any system of measures must reflect the *efficiency and productivity* of employees, the Committee expects that the IRS will establish a *balanced system of measures* that will ensure that taxpayer satisfaction is paramount throughout all IRS functions.”) (Emphasis added).

## ANALYSIS OF PROBLEM

### Background

#### *The IRS is required to use measures to evaluate its programs.*

The IRS is legally required to establish performance targets and annually report its progress in meeting them.<sup>3</sup> In addition, the Commissioner or Deputy Commissioner meets with Division Commissioners and Chiefs to discuss program-level measures as part of a quarterly Business Performance Review (BPR) process.<sup>4</sup> Many of the same measures also appear in the individual performance appraisals of IRS executives.<sup>5</sup>

#### *It is difficult for the IRS to measure results without creating counterproductive incentives.*

About 97 percent of all tax revenues the IRS collects come in without the need for direct enforcement actions – because of voluntary compliance.<sup>6</sup> However, it has always been easier to quantify and measure enforcement actions and direct enforcement revenue than the indirect effect of IRS activities – including taxpayer service – on voluntary compliance.

In the years before 1998, the IRS's enforcement-oriented measures created incentives for enforcement behavior that alienated the public without enhancing tax compliance.<sup>7</sup> The National Commission on Restructuring the IRS found that measures of short-term performance and efficiency did not align with the IRS's mission.<sup>8</sup>

#### *The IRS Restructuring and Reform Act of 1998 (RRA 98) required the IRS to adopt “balanced measures” to counter “efficiency and productivity” metrics.*

RRA 98 prohibited the IRS from rating employees based on “records of tax enforcement results” (ROTERS).<sup>9</sup> It also required the IRS to establish a new system to measure both individual and group performance, which was supposed to focus on service.<sup>10</sup> Legislators expected new customer satisfaction measures to counter the negative incentives created by short-term measures of efficiency and productivity.<sup>11</sup> Similarly, policymakers proposed

<sup>3</sup> See 31 U.S.C. § 1115. See also Department of the Treasury, Annual Performance Report, FY 2009, 133-34 (Feb. 1, 2010), available at [http://www.ustreas.gov/offices/management/dco/accountability-reports/09Treas\\_APR\\_full%20book\\_05.pdf](http://www.ustreas.gov/offices/management/dco/accountability-reports/09Treas_APR_full%20book_05.pdf) (including the IRS measures).

<sup>4</sup> IRM 1.5.1.15 (Aug. 21, 2009). The IRS Oversight Board may also review the BPRs.

<sup>5</sup> See, e.g., IRM 1.5.2.20 (June 1, 2010).

<sup>6</sup> See IRS, *Tax Gap Map for Year 2001* (Feb. 2007), available at [http://www.irs.gov/pub/irs-utl/tax\\_gap\\_update\\_070212.pdf](http://www.irs.gov/pub/irs-utl/tax_gap_update_070212.pdf) (estimating that of the \$1.822 trillion that would be collected for 2001, 97 percent (\$1.767 trillion) would come in timely and voluntarily, and only three percent (\$55 billion) would come in as a result of enforcement; another \$290 billion – the so-called net tax gap – would remain uncollected).

<sup>7</sup> Commissioner Charles O. Rossotti, *Report to the IRS Oversight Board Assessment of the IRS and the Tax System* 3 (Sept. 2002).

<sup>8</sup> Report of the National Commission on Restructuring the IRS, *A Vision for a New IRS* 22 (June 25, 1997).

<sup>9</sup> See RRA 98, Pub. L. No. 105-206, 112 Stat. 722, Title I, §§ 1204, 9508(a)(2) (1998); Joint Committee on Taxation (JCT), JCS-6-98, *General Explanation of Tax Legislation Enacted in 1998* 47-50 (1998). See also IRM 1.5.2.2 (June 1, 2010) (noting that both the IRS and Congress had banned the use of ROTERS in rating front line employees before 1998, but the bans were “not necessarily easy to understand and were not interpreted consistently”). The number of liens, levies, or dollars collected or assessed are examples of ROTERS. IRM 1.5.2.6 (June 1, 2010).

<sup>10</sup> RRA 98, Pub. L. No. 105-206, 112 Stat. 722, Title I, § 9508(a)(2) (1998). See also Report of the National Commission on Restructuring the IRS, *A Vision for a New IRS* 5 (June 25, 1997).

<sup>11</sup> IRS Restructuring and Reform Act of 1997, H.R. Rept. No. 105-364, at 47 (1997).

new employee satisfaction measures that would reduce the incentive for IRS managers to apply indirect pressure on employees to meet organizational productivity and efficiency targets.<sup>12</sup> Legislators believed these changes would improve voluntary compliance by creating incentives for activities that would foster trust in the IRS.<sup>13</sup>

Policymakers borrowed this “balanced measures” approach from the private sector where companies use customer satisfaction measures to counter an over-emphasis on short-term financial results at the expense of long-term customer relationships.<sup>14</sup> Accordingly, the IRS established three “balanced measures:” customer satisfaction, employee satisfaction, and business results.<sup>15</sup>

### *Which “efficiency and productivity” metrics need to be counter-balanced?*

Cycle time is an efficiency or productivity metric, but true measures of timeliness are not. Cycle time measures generally reflect the speed with which an IRS program closes cases or takes other actions. By contrast, timeliness measures indicate whether the IRS was timely in taking a discrete action from the taxpayer’s perspective (*e.g.*, did the IRS return the taxpayer’s call within three days) without regard to how many actions it took or cases it closed during the reporting period. While timeliness measures provide an incentive to move cases along by taking specific actions in a timely manner, they do not necessarily create incentives to close cases prematurely, as efficiency and productivity measures can.<sup>16</sup>

### *The IRS planned to reevaluate the incentives its measures create.*

Following RRA 98, IRS measures continued to evolve, and the current IRS Commissioner pledged to reexamine the unintended incentives they create.<sup>17</sup> We intend this discussion to assist the Commissioner’s review.<sup>18</sup>

<sup>12</sup> See Report of the National Commission on Restructuring the IRS, *A Vision for a New IRS* 22 (June 25, 1997); T.D. 8830, 64 Fed. Reg. ¶ 42,834 (Aug. 6, 1999) (explaining “[t]he presence of measures that evaluate the quality of the work done by the unit, the satisfaction of customers served by the unit (including taxpayers), and the satisfaction of employees working in the unit will obviate the risk that managers place undue emphasis upon the quantity of work completed.”).

<sup>13</sup> JCT, JCS-6-98, *General Explanation of Tax Legislation Enacted in 1998*, 19 (1998) (noting that RRA 98 was intended to restore public confidence in the IRS, in part, because “the Congress believed that most Americans are willing to pay their fair share of taxes, and that public confidence in the IRS is key to maintaining that willingness.”).

<sup>14</sup> See, *e.g.*, Government Accountability Office (GAO), GAO-01-234, *IRS Should Enhance Its Performance Management System* 6 (Feb. 2001); Charles O. Rossotti, *Many Unhappy Returns: One Man’s Quest to Turn Around the Most Unpopular Organization in America* 112-115 (Harvard Business School Press 2005).

<sup>15</sup> 26 C.F.R. § 801.1 *et seq.* See also IRM 1.5.1 (Aug. 21, 2009).

<sup>16</sup> The IRS has issued confusing guidance about whether “cycle time” and “overage” are quality or quantity measures. Compare 26 C.F.R. § 801.6(c) (defining “cycle time” and number or percentage of “overage” cases as quantity measures) and IRM 1.5.2.8 (June 1, 2010) (same), with IRM 1.5.2.9 (June 1, 2010) (defining “overage cases” and “cycle time” as quality measures).

<sup>17</sup> Responses to Questions Following the Testimony of Doug Shulman, Nominee for IRS Commissioner, Before the Senate Committee on Finance, Q/A 45 (Feb. 11, 2008) (pledging to review IRS measures to “make sure that they are aligned with the priorities of the IRS as a whole... are clear, and ... engender the desired behavior.”).

<sup>18</sup> For prior coverage of the impact of IRS measures, see, *e.g.*, National Taxpayer Advocate 2008 Annual Report to Congress 159.

**IRS program-level measures focus on processing speed.***IRS leaders focus on “business results” when evaluating programs.*

The IRS generally measures customer and employee satisfaction using annual (or sometimes quarterly) surveys.<sup>19</sup> However, the IRS has stated that they are lagging indicators that can be difficult to link to specific activities or programs.<sup>20</sup> Moreover, SB/SE only includes program-level customer satisfaction results in its BPR on an annual basis.<sup>21</sup> Thus, it may be natural for IRS leaders to focus primarily on quarterly “business result” measures, particularly those that regularly appear in the BPR or other highly visible documents.

*By design, business-result measures focus on cycle time.*

The IRS is supposed to assess operational business results using measures of “output or quantity” and “efficiency or quality.”<sup>22</sup> Each of these may reflect cycle time, however, because output or quantity is production over a reporting period, and “efficiency” is simply quantity divided by another statistic such as the number of full time equivalent (FTE) employees. Even measures of “quality” can include measures of cycle time (rather than timeliness, as discussed above).<sup>23</sup> In other words, an improvement (decrease) in cycle time can improve all of the business result measures. Thus, notwithstanding RRA 98’s direction for the IRS to focus on customer satisfaction, by design, processing speed (or cycle time) is still the most important measure.

*In practice, the BPR focuses on cycle time.*

The only measures referenced in SB/SE’s discussion of the “performance” of each of its five major programs in the BPR is a table that lists productivity, cycle time, efficiency, or overage.<sup>24</sup> For example, the “collection performance” section of SB/SE’s BPR reports only:

- The number of liens, levies, seizures, unfiled return case closures (called “taxpayer delinquency investigations” (TDI)), and unpaid tax case closures (called “taxpayer delinquent accounts” (TDA));

<sup>19</sup> IRM 1.5.1.7 (Aug. 21, 2009); IRM 1.5.1.8 (Aug. 21, 2009); IRM 1.5.1.9 (Aug. 21, 2009); IRM Exhibit 1.5.1-1 (Q5), (Q6) (Aug. 21, 2009); Wage and Investment (W&I) Division, Business Performance Review (BPR) 29-30 (Feb. 2010).

<sup>20</sup> As one recent W&I Division BPR explained:

In Reporting Compliance, the success of previous customer satisfaction improvement projects have [sic] often been measured only by the change in our customer satisfaction scores, even though that measure is a “lagging indicator” at best and is often affected by multiple other variables. We would like to look at adding other metrics for success when we do our projects. W&I, Business Performance Review 29-30 (Feb. 2010).

<sup>21</sup> IRS Response to TAS information request (Oct. 27, 2010); SB/SE, Business Performance Review (Feb. 2010) (not listing program-level survey results); SB/SE, Business Performance Review (May 2010) (not listing program-level survey results). In its fourth quarter 2010 BPR, SB/SE plans to explain that the only significant change to customer satisfaction results – a decline in satisfaction with field examinations – is attributable to “changes to the survey and the mix of work.” IRS Response to TAS information request (Oct. 27, 2010). It may be more difficult to explain why none of its initiatives have had a significant positive impact on customer satisfaction results. *Id.*

<sup>22</sup> IRM 1.5.1.8.2 (Aug. 21, 2009). The IRM seems to equate quality with efficiency.

<sup>23</sup> IRM 1.5.1.8.2 (Aug. 21, 2009); IRM Exhibit 1.5.1-1 (Q15) (Aug. 21, 2009) (explaining that cycle time and overage are considered “quality” measures (rather than quantity measures) because they “assess the timeliness in which work is completed [and] do not tell us how much work was competed.”); IRM 1.5.2.9 (June 1, 2010).

<sup>24</sup> See SB/SE, Business Performance Review (May 2010). The five programs include Campus Compliance Services, Fraud/Bank Secrecy Act, Collection, Specialty Programs, and Examination. *Id.*

- The percentage of TDI/TDA cases that are overage (*i.e.*, have been open for 16 months or longer); and
- The percent of field offers in compromise (OIC) closed within nine months, along with an indication of whether Collection is on target to meet its production goals.<sup>25</sup>

Thus, all of the measures the IRS lists in the BPR to assess the success or failure of the Collection Field function improve if the IRS increases processing speed.<sup>26</sup> The IRS does not include any other measures to provide balance (*e.g.*, employee satisfaction, taxpayer satisfaction, future compliance, or delinquencies resolved without the need for a lien, levy, etc.).<sup>27</sup>

SB/SE asserts that it also uses an expanded suite of measures to evaluate the Collection Field function, as shown in the following table.<sup>28</sup>

<sup>25</sup> See SB/SE, Business Performance Review (May 2010); IRM Exhibit 3.13.2-1 (Jan. 1, 2010) (defining TDI and TDA). The IRS has targets for TDA Closures, TDI Closures, Percent Overage – TDA/TDI Taxpayer Combo, and Percent OIC Field Closures in 0-9 Months. SB/SE, Business Performance Review (May 2010).

<sup>26</sup> SB/SE's BPR generally describes actions designed to improve customer and employee satisfaction, but these descriptions are not performance measures.

<sup>27</sup> Given the IRS's focus on cycle time, it is not surprising that some IRS employees feel pressure to increase the production figures reflected in the BPR. The 2010 employee survey results show that, notwithstanding RRA 98's prohibition on ROTERS, 17.2 percent of IRS enforcement employees indicated that "enforcement results were verbally communicated as production goals or quotas by either my supervisor or next-higher level manager;" and 9.1 percent indicated that "tax enforcement results were used by either my supervisor or next-higher level manager to evaluate my performance." Sirota Survey Intelligence Data (Sept. 17, 2010).

<sup>28</sup> We cannot verify how SB/SE uses these measures because SB/SE declined to provide TAS with the documents in which they appear. SB/SE response to TAS information request (Aug. 19, 2010).

**TABLE 1.3.1, Incentives Produced by Other Collection Performance Measures and Indicators<sup>29</sup>**

Measure/Indicator	Incentive			
	Productivity, speed, or efficiency	Timeliness	Accuracy, completeness, or quality	Effect on taxpayers or employees
Number of Taxpayer Closures	X			
Taxpayer Closures per FTE	X			
\$ Collected From TDAs Assigned to Field	X			
Number of Total \$ Collected	X			
Number of TDA Closures - TDA & Combo Taxpayers	X			
Number of TDA Closures - Module	X			
TDA & Combo taxpayers per FTE	X			
Cycle Time TDA Module - Avg. Weeks	X			
Percent Overage - TDA/Combo TDI Taxpayers Combined <sup>30</sup>		X		
Number Overage-TDA/Combo TDI Taxpayers Combined		X		
Number of Liens	X			
Number of Levies	X			
Number of Seizures	X			
Number of TDI Closures - Stand Alone Taxpayers	X			
Number of TDI Closures - Modules	X			
TDI Stand Alone taxpayers per FTE	X			
Cycle Time TDI Total Investigations - Avg. Weeks	X			
Number of OIC Field Closures	X			
Number of OIC Field Receipts				X
Percent OIC Field Closures in 0 - 9 Months	X			
Employee Satisfaction (Percent Satisfied)				X
Taxpayer Satisfaction (Percent Satisfied)				X
Taxpayer Satisfaction (Percent Dissatisfied)				X
NQRS Quality Score			X	
NQRS OIC Quality Score			X	
<b>Number</b>	<b>17</b>	<b>2</b>	<b>2</b>	<b>4</b>
<b>Percentage</b>	<b>68%</b>	<b>8%</b>	<b>8%</b>	<b>16%</b>

While this expanded list includes measures of timeliness, accuracy, customer satisfaction, and employee satisfaction, it focuses primarily on measures of productivity, speed, and

<sup>29</sup> SB/SE response to TAS information request (Aug. 19, 2010). This table omits two staffing measures: “FTEs - Total FAC 7D plus FAC 7P” and “FTEs - Areas.” We omitted them because they are unlikely to create the kind of incentives that we are discussing here. For purposes of the tables included in this discussion, measures included in the “productivity, speed, or efficiency” column reflect the number of actions, such as case closures or dollars collected during the reporting period, whether or not presented as a ratio with another statistic. By contrast, measures included in the “timeliness” column reflect how timely the IRS was in taking a discrete action, such as contacting a taxpayer, without regard to how many actions it took or cases it closed during the period.

<sup>30</sup> “Overage” TDA/TDI means “taxpayers in ending inventory that have been in progress for 16 months or longer.” SB/SE, Field Collection, FY 2010 Data Dictionary, Business Performance Review Measures (Dec. 2009). While measures of overage inventory have the potential to create incentives similar to productivity measures, for purposes of this analysis, we conservatively include them in the timeliness column because they do not necessarily create an incentive for the IRS to reduce cycle time or to close cases prematurely, particularly if the overage benchmark is not excessively short.

efficiency, which represent over two-thirds of the total. This is because SB/SE breaks out productivity, speed, and efficiency for every type of collection activity separately (*e.g.*, liens, levies, seizures, TDA/TDI closures), but does not similarly break out timeliness, accuracy, completeness, or satisfaction.

*Omitting satisfaction and quality measures from the BPR reduces balance.*

Disaggregating or separating processes and measures that improve with cycle time (*i.e.*, productivity and efficiency measures), but not doing so for other measures reduces the effectiveness of the balanced measures concept. Moreover, readers of SB/SE's quarterly BPR or other high-level collection reports have no way to know:

- If Collection is processing liens, levies, seizures, offers, TDI, and TDA cases in a way that enhances future compliance;
- If Collection is applying its resources to the right cases to maximize voluntary compliance; or
- How long Collection is allowing delinquencies to remain unaddressed, or only partially or temporarily addressed, from the taxpayer's (and the government's) perspective.

The IRS assumes that an increase in the number (and speed) of quarterly collection activities is the primary way to achieve the IRS mission to improve voluntary compliance.<sup>31</sup> However, the IRS has not done the research necessary to validate this conclusion, and both the IRS and legislators previously rejected it when the balanced measurement system was adopted.<sup>32</sup>

The IRS's focus on processing speed is not due to a lack of other measures. Collection tracks other items that may be better measures of long-term success in promoting voluntary compliance. For example, it tracks the default rate for installment agreements and offers in compromise, the rate at which taxpayers cure delinquencies after receiving an IRS

<sup>31</sup> According to the IRS, "we can assess how we are doing on Improving Voluntary Compliance (is it going up or going down) through the budget level measures reported for Examination and Collection coverage." IRS response to TAS information request (Sept. 24, 2010). The Office of Management and Budget (OMB) website similarly states:

IRS measures the effects of collection program activities on the VCR [voluntary compliance rate] (which includes filing and payment compliance) through the long-term proxy measures of Collection Coverage and Collection Efficiency....the effects of the Collection program on taxpayer behavior is also evidenced through deterrence, though the effect as such is not measured by the IRS.

OMB, *Detailed Information on the Internal Revenue Service Tax Collection Assessment* (2008), <http://www.whitehouse.gov/omb/expectmore/detail/10000424.2008.html> (last updated Jan. 9, 2009).

<sup>32</sup> Research suggests that both enforcement and service activities have an effect on voluntary compliance. See, *e.g.*, Alan H. Plumley, Pub. 1916, *The Determinants of Individual Income Tax Compliance: Estimating the Impacts of Tax Policy, Enforcement, and IRS Responsiveness* 35-36 (Oct. 1996). But, not all enforcement activities have the same effect. *Id.* Some may reduce compliance. See National Taxpayer Advocate 2009 Annual Report to Congress vol. 2, 29 (finding about 74 percent of those taxpayers with no assessed delinquency at the end of 2001 who did not pay a 2002 assessment, had a subsequent-year delinquency or unfiled return).

notice, and the rate at which delinquencies are resolved without the need for enforcement.<sup>33</sup> However, Collection does not highlight such data as part of the formal BPR process.<sup>34</sup>

*Enterprise-level reports also focus on speed.*

The IRS places a similar emphasis on processing speed at the enterprise level. Nearly half of the IRS’s enterprise-level measures reflected in its Annual Performance Report (called “budget level” measures) provide incentives for productivity, speed, and efficiency, as shown in the following table.

**TABLE 1.3.2, Incentives Produced by IRS Enterprise-Level Measures and Indicators<sup>35</sup>**

Measure/Indicator	Incentive			
	Productivity, speed, or efficiency	Timeliness	Accuracy, completeness, or quality	Effect on taxpayers or employees
<b>Goal 1: Improve Service to Make Voluntary Compliance Easier</b>				
Customer Service Representative (CSR) Level of Service	X			
Customer Contacts Resolved per Staff Year	X			
Customer Accuracy – Tax Law Phones			X	
Customer Accuracy – Customer Accounts (Phones)			X	
Timeliness of Critical Individual Filing Season Tax Products to the Public		X		
Timeliness of Critical Tax Exempt and Government Entities (TE/GE) and Business Tax Products to the Public		X		
Percent Individual Returns Processed Electronically				X
Cost per Taxpayer Served (\$) (HCTC)	X			
Sign-Up Time (Days) – Customer Engagement (HCTC)				X
Percent Business Returns Processed Electronically				X
Refund Timeliness – Individual (Paper)		X		
Taxpayer Self Assistance Rate				X
Service Subtotal – Number	3	3	2	4
Service Subtotal – Percent	25%	25%	17%	33%
<b>Goal 2: Enforcement the Law to Ensure Everyone Meets Their Obligation to Pay Taxes</b>				
Examination Coverage – Individual	X			
Field Examination National Quality Review Score			X	
Office Examination National Quality Review Score			X	
Examination Quality – Industry			X	
Examination Quality – Coordinated Industry			X	
Examination Coverage – Business (Corps. >\$10M)	X			

table continued on next page

<sup>33</sup> SB/SE response to TAS information request (Aug. 2, 2010).

<sup>34</sup> According to SB/SE, it is “in the process of establishing objectives and performance measures for key collection notices. Data is available and tracked for the number of case dispositions in notice status.” SB/SE response to TAS information request (Aug. 2, 2010).

<sup>35</sup> See Department of the Treasury, Annual Performance Report, FY 2009, 133-34 (Feb. 1, 2010). This table omits two Business System Modernization measures because they have little impact on incentives facing IRS programs that serve taxpayers.

Measure/Indicator	Incentive			
	Productivity, speed, or efficiency	Timeliness	Accuracy, completeness, or quality	Effect on taxpayers or employees
Examination Efficiency - Individual (1040)	X			
Automated Underreporter (AUR) Efficiency	X			
Automated Underreporter (AUR) Coverage	X			
Collection Coverage - Units	X			
Collection Efficiency - Units	X			
Field Collection Embedded Quality			X	
Automated Collection System (ACS) Accuracy			X	
Criminal Investigations Completed	X			
Number of Convictions	X			
Conviction Rate			X	
Conviction Efficiency Rate (\$)	X			
TE/GE Determination Case Closures	X			
Enforcement Subtotal - Number	11	0	7	0
Enforcement Subtotal - Percent	61%	0%	39%	0%
<b>Total - Number</b>	<b>14</b>	<b>3</b>	<b>9</b>	<b>4</b>
<b>Total - Percent</b>	<b>47%</b>	<b>10%</b>	<b>30%</b>	<b>13%</b>

While only about 25 percent of the service measures provided an incentive to reduce processing time, more than 60 percent of the IRS’s enforcement metrics do so.

In addition, none of the IRS’s enforcement measures reflect the timeliness of the enforcement from the taxpayer’s perspective or its impact on compliance. It is possible to measure how long a taxpayer was delinquent before the IRS fully addressed the problem. It is also possible to measure the impact of the IRS on a taxpayer’s future compliance, but the IRS does not do so – at least not for purposes of the highly visible Annual Performance Report or the BPRs described above.

**The IRS’s inability to measure service revenues combines with budget rules to provide incentives for IRS programs to focus on enforcement revenue.**

Enforcement initiatives that will generate direct enforcement revenue receive disproportionate funding compared to other enforcement or service initiatives that could bring in larger – but difficult to measure – revenues by enhancing voluntary compliance. Pursuant to a budget rule called the “program integrity cap,” if the IRS estimates the return on investment (ROI) for an initiative is greater than one, Congress can appropriate funds for it without cutting spending elsewhere.<sup>36</sup> ROI is the ratio of the direct revenue the IRS expects an initiative to generate divided by the direct and indirect costs of that initiative (*e.g.*, where direct revenue is \$3x and total costs are \$1x, the ROI is 3 or 3:1). To date, Congress has

<sup>36</sup> For prior discussion of the budget rules, see National Taxpayer Advocate 2006 Annual Report to Congress 442-446.

only used this budget device to fund enforcement.<sup>37</sup> Congress has not been able to use it to fund services, in part, because the IRS does not estimate the ROI for service initiatives.

However, in some cases the IRS's estimates of enforcement revenue are just as speculative as any estimate of service revenue would be. The IRS cannot reliably project the amount of controversy an enforcement initiative will generate – particularly a new initiative – and does not even try to estimate the impact of initiatives on voluntary compliance (*i.e.*, indirect revenue gains or losses).<sup>38</sup> For example, to the extent taxpayers consider an initiative unfair (*e.g.*, because it is overly automated or rushed as a result of a focus on cycle time), it could potentially generate more TAS cases, Appeals cases, and litigation than expected. In such instances, the initiative could also reduce voluntary compliance and indirect revenue, which the IRS does not attempt to measure.<sup>39</sup> Consequently, the IRS could project an initiative will produce a positive ROI, when in fact the ROI is negative.

In addition, as noted above, about 97 percent of all tax revenues the IRS collects come in without the need for expensive enforcement resources, in part because of IRS services.<sup>40</sup> If the IRS did not process submissions; provide forms, instructions, and guidance; send notices and communicate with taxpayers and practitioners in other ways; and make taxpayers feel the IRS and its processes are fair, it would not collect these “service revenues.”<sup>41</sup> Yet, Congress cannot use the program integrity cap to fund services, in part, because the IRS does not measure or quantify the ROI for taxpayer services.

Predictably, funding for enforcement increased more than five times as much as funding for services between fiscal year (FY) 2007 and FY 2010.<sup>42</sup> Thus, notwithstanding RRA 98's direction for the IRS to become more service-oriented and the increasing number of social benefit programs (including Economic Stimulus Payments, the Making Work Pay credit, the First-Time Homebuyer Credit, and a variety of provisions relating to health care reform) that Congress is asking the IRS to administer, it is becoming less of a service organization and more of an enforcement agency.

Moreover, if Congress funds IRS programs likely to produce the most direct enforcement revenue but not those that produce service revenue or improve voluntary compliance, the

<sup>37</sup> See, *e.g.*, Department of the Treasury, *Budget in Brief FY 2011*, 2, 10 (2010), <http://www.irs.gov/pub/newsroom/budget-in-brief-2011.pdf>.

<sup>38</sup> Each IRS business unit does attempt to estimate how many FTEs will be required directly and indirectly to handle the down stream consequences of the initiative. IRS response to TAS information request (Aug. 6, 2010).

<sup>39</sup> See, *e.g.*, Swedish Tax Agency, *Right From The Start, Research and Strategies* 6-7, (Aug. 2005) (surveying research suggesting that people are less likely to accept a negative decision or voluntarily comply if they do not perceive the tax agency to be fair).

<sup>40</sup> See IRS, *Tax Gap Map for Year 2001* (Feb. 2007).

<sup>41</sup> However, the IRS recently proposed to save money by discontinuing tax form mailings. IRS, FY 2011 Congressional Budget Submission 19 (Feb. 1, 2010). It recently discontinued these mailings. See IRS, *Tax Package Mailing to End Following Growth of e-File*, available at <http://www.irs.gov/businesses/small/article/0,,id=228152,00.html> (last visited Oct. 12, 2010).

<sup>42</sup> Nominal funding for services increased by \$141 million or 6.6 percent (from \$2.138 billion in FY 2007 to \$2.279 billion in FY 2010), while nominal funding for enforcement increased by \$818 million or 17.5 percent (from \$4.686 billion to \$5.504 billion) over the same period. GAO, GAO-10-687R, *Assessment of Budget Justification for Fiscal Year 2011 Identified Opportunities to Enhance Transparency* 10 (May 2010). We divided the increase for enforcement (\$818 million) by the increase for services (\$141 million) to compute that the increase for enforcement was 5.8 times the increase for services.

implied message to IRS leaders is that they should care more about the impact of a program on direct enforcement revenue than on voluntary compliance. The IRS reinforces this message with measures that focus primarily on cycle time rather than customer satisfaction, accuracy, completeness, timeliness or other factors likely to increase voluntary compliance.

**A focus on cycle time measures and enforcement revenue leads to reactive (rather than proactive) policies and procedures.**

The IRS's focus on quick closures and enforcement revenue provides incentives for IRS leaders to adopt policies and procedures that result in premature IRS "closures" that do not necessarily further the government's long-term interests or fully resolve the taxpayer's problems.<sup>43</sup> For example, in last year's report TAS discussed how IRS collection executives issued guidance requiring employees to obtain approval not to file a notice of federal tax lien (NFTL) even in cases where the lien would not attach to anything or would impair the taxpayer's ability to earn income needed to repay the government.<sup>44</sup> This policy is likely to increase the NFTL numbers shown on the BPR along with enforcement revenue statistics, but will not necessarily resolve taxpayer delinquencies more quickly, increase government revenue, or improve future compliance.<sup>45</sup> A focus on cycle time and cost also provided an incentive for the IRS to expand its use of correspondence examinations to more complex business returns, even though the procedure's inability to detect unreported income on such returns may actually reduce voluntary compliance.<sup>46</sup> If the IRS's program-level measures focused more on the specific and measurable impact of IRS activities on taxpayers (*e.g.*, satisfaction and future compliance), they would create a stronger incentive for IRS leaders to adopt policies that achieve long-term results and ultimately bring in more revenue.

## CONCLUSION

The IRS's organizational measures provide incentives for leaders to promulgate policies that maximize processing speed. Similarly, the IRS's ability to quantify the ROI for enforcement (and thereby obtain funding for it) coupled with its inability to do the same for taxpayer service creates an incentive for IRS leaders to adopt policies that *generate* direct enforcement revenue quickly but not those that *prevent* delinquencies. Such policies are likely to send taxpayers from one program to another without necessarily reducing IRS-wide cycle time, preventing, addressing or resolving taxpayer problems, or promoting

<sup>43</sup> For additional discussion of IRS measures, see National Taxpayer Advocate 2008, Annual Report to Congress 159.

<sup>44</sup> SB/SE, No. 05-1208-069, *Interim Guidance for Approval of Lien Determinations* (Dec. 22, 2008).

<sup>45</sup> For a more detailed discussion, see National Taxpayer Advocate 2009 Annual Report to Congress 17-40.

<sup>46</sup> See, *e.g.*, SB/SE, Campus Compliance Services Program Letter for FY 2010 (Sept. 28, 2009) (expressing the goal to "expand workload selection in Correspondence Examination to complex Business Master File (BMF) and Individual Master File (IMF) workloads"); National Taxpayer Advocate 2009 Annual Report to Congress 158, 166-67. Similarly, a focus on minimizing the time IRS employees spend on calls provides an incentive for policies that end communications prematurely. See, *e.g.*, IRM 1.4.16.3.4.2 (Jan. 1, 2009); W&I, Business Performance Review 17 (Feb. 2010) (listing measures of calls answered and resolved, but no measures of how often taxpayers required additional communications that should not have been necessary).

voluntary compliance. Thus, the IRS could do a better job of aligning program-level incentives with its underlying mission to promote voluntary compliance.

The National Taxpayer Advocate offers these preliminary recommendations to the IRS:

1. Report disaggregated satisfaction, accuracy, and completeness metrics at every level at which the IRS reports efficiency, speed, or productivity metrics;
2. Create measures for the BPR and other highly visible reports that reflect the period between the due date of the return and final resolution of any liability (*e.g.*, through full payment, abatement, or compromise), as previously recommended by the National Taxpayer Advocate;<sup>47</sup>
3. Create measures for the BPR and other highly visible reports that reflect the extent to which the IRS's programs completely resolved the noncompliance and prompted the taxpayer to comply in subsequent periods; and
4. Initiate research to quantify the impact of various types of taxpayer service and outreach on voluntary compliance and direct tax revenue so the IRS can estimate the ROI of such initiatives.<sup>48</sup>

### IRS COMMENTS

The IRS Restructuring and Reform Act of 1998 initiated sweeping changes that had a dramatic effect on IRS operations. The most visible change was the creation of four business operating divisions, each with end-to-end accountability for a given customer segment. This organizational structure enabled IRS leadership to gain a deeper understanding of the IRS's customer segments and take a more holistic view of their needs. The IRS also adopted a new mission statement, which captured its twin purposes to “[1] provide America's taxpayers top-quality service by helping them understand and meet their tax responsibilities and [2] enforce the law with integrity and fairness to all.”<sup>49</sup> This dual mission recognizes that voluntary compliance depends on a combination of high-quality service delivery and a credible enforcement presence. A final significant change was the introduction of the balanced measures approach encompassing customer satisfaction, employee engagement, and business results, which remains in place today.<sup>50</sup> As a consequence of these changes, we believe that the IRS today effectively balances its twin missions of service *and* enforcement to promote voluntary compliance and ensure that everyone pays their fair share.

The National Taxpayer Advocate contends that “IRS program-level measures focus on processing speed.” As evidence, they focus on the quarterly SB/SE Business Performance

<sup>47</sup> See, *e.g.*, National Taxpayer Advocate 2006 Annual Report to Congress 62, 81.

<sup>48</sup> The National Taxpayer Advocate previously made similar recommendations. See, *e.g.*, National Taxpayer Advocate 2007 Annual Report to Congress 35, 38; National Taxpayer Advocate 2006 Annual Report to Congress 442-446.

<sup>49</sup> IRS, *The Agency, Its Mission and Statutory Authority*, <http://www.irs.gov/irs/article/0,,id=98141,00.html> (last reviewed or updated, June 23, 2010) (numbering added).

<sup>50</sup> See IRM 1.5.1.

Review, and specifically, the Collection Field function. However, the quarterly BPR is a high-level document that cannot reasonably contain all the measures used to manage operations. SB/SE leadership, for instance, regularly reviews additional performance data beyond the measures shown in BPR, including the 25 performance indicators for Collection shown in Table 1.3.1, *Incentives Produced by Collection Performance Measures and Indicators*. While this is only a subset of the full range of operational measures that Collection management uses to monitor and assess its programs,<sup>51</sup> this list contains the key elements of a balanced measures system—customer satisfaction, employee engagement and business results (both quality and quantity). However, based on this list, the National Taxpayer Advocate concludes that 68 percent of Collection measures incentivize “productivity, speed, or efficiency.”

This analysis presumes that a given measure cannot serve multiple purposes. In particular, the classification of cycle time as a measure of speed is incomplete. Clearly, cycle time is also closely related to “timeliness,” as maintaining timely interactions throughout the collection process is an essential component of delivering cycle time goals. Moreover, timeliness/cycle time is a significant customer concern. The “amount of time you had to spend on this collection issue” was the top improvement priority for collection customers on the most recent Collection customer satisfaction survey.<sup>52</sup> This question had the highest level of dissatisfaction (25 percent) on the survey. Clearly, customers share with the IRS the desire to reduce the length of the process and resolve their cases efficiently.

Other measures similarly create multiple incentives. Like cycle time, overage measures have a direct effect on taxpayers eager to resolve their IRS issues (and also incentivize speed or efficiency). Quality scores are based on an evaluation of multiple quality attributes, including timely interactions and customer relations as well as accuracy.<sup>53</sup> If these overlapping incentives were captured, a more balanced picture of Collection’s suite of measures would emerge. Even then, the mere number of measures is not an indicator of their relative significance. Many of the “productivity, speed, or efficiency” measures are more appropriately viewed as a set (*e.g.*, *Number of TDA Closures – TDA & Combo Taxpayers*, *Number of TDA Closures – Module*, *Number of TDI Closures – Stand Alone Taxpayers*, and *Number of TDI Closures – Module*). Such measures capture distinct aspects of the workload for operational purposes, but tend to be considered collectively when evaluating the program overall. Other performance indicators (in particular, measures of dollars collected, and the numbers of liens, levies and seizures) are reported solely for informational purposes but do not have targets established and are not used to evaluate performance.

<sup>51</sup> The Enterprise Collection Report, for example, tracks the overall outcome of dispositions by monitoring rework (*e.g.*, installment agreements or offers in compromise that are returned to inventory) to help measure the success of Collection dispositions. The monthly Area Assessment Report breaks down details behind the overall quality score. Quality measures are also routinely analyzed to determine if a correlation exists between a particular quality element and customer satisfaction.

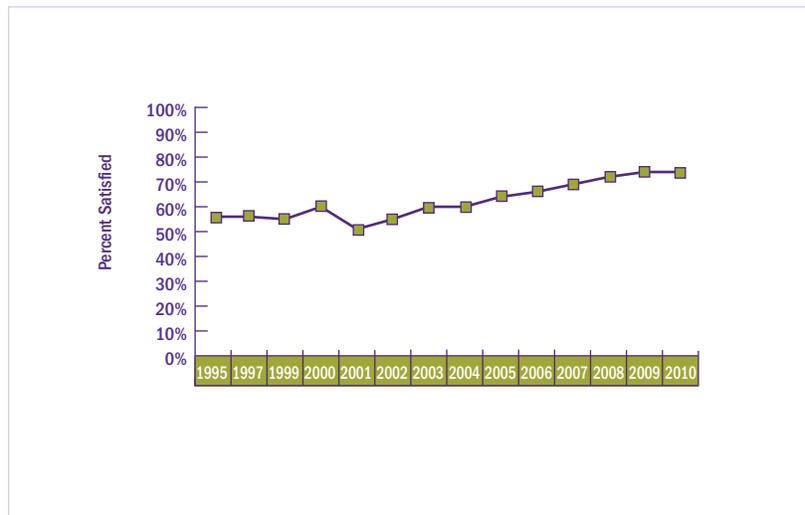
<sup>52</sup> Pacific Consulting Group, *SB/SE Field Collection National Report (April – June 2010)*.

<sup>53</sup> Relevant attributes include: *Timely Employee Actions*, *Timely Follow-Up Actions*, *Clear Action Dates and Expectations*, *Advised [Taxpayer] of Consequences*, and *Requested/Secured Financial Information*, among others. (Source: Field Collection Embedded Quality Manager and Reviewer Job Aid).

Applying equivalent weights to these metrics to determine the percentage associated with each incentive category is potentially misleading as this approach does not reflect how they are actually used to manage the Collection program.

The National Taxpayer Advocate conducts a similar analysis of the IRS enterprise-level budget measures reported in the IRS’s Annual Performance Report, concluding that “IRS places a similar emphasis on processing speed at the enterprise level.” Again, however, this source includes only a small subset of the many measures used to monitor and evaluate operational performance. The 30 measures in the Annual Performance Report were selected in conjunction with Treasury to provide a high-level perspective on each IRS budget activity. By contrast, the quarterly BPRs for the four IRS business operating divisions contain over 300 measures, ranging from closures and cycle time to quality and customer satisfaction. Moreover, paralleling the budget-level measures, IRS reports a separate set of enterprise-level measures each quarter to the IRS Oversight Board.<sup>54</sup> The measures monitored by the Board focus solely on cycle time and customer satisfaction, reflecting the Board’s perspective that cycle time is a key driver of customer satisfaction.<sup>55</sup>

**FIGURE 1.3.3, IRS Employee Satisfaction**



Ultimately, though, the significance of IRS program measures lies not in what measures are reported to which level, but in the impact the balanced measures approach has had on the management and operation of the IRS. To address employee engagement, IRS has conducted an all-employee survey since at least 1995 (annually since 1999). Following a decline in 2001 during the agency reorganization, employee satisfaction has risen steadily through the

<sup>54</sup> The IRS Oversight Board was created by Congress under the IRS Restructuring and Reform Act of 1998 with responsibility to oversee the IRS in its administration, management, conduct, direction, and supervision of the execution and application of the internal revenue laws.

<sup>55</sup> The 16 “enforcement” measures tracked by the Oversight Board are comprised exclusively of cycle time and customer satisfaction measures. The Board also monitors key measures of IRS service and support functions.

current year, as shown in Figure X, *IRS Employee Satisfaction*.<sup>56</sup> Meanwhile, the IRS score on the Partnership for Public Service's *Best Places to Work in the Federal Government* index has climbed from 56.1 in 2003 to 68.2 in 2010.<sup>57</sup> While our rank of 79 out of 224 agency subcomponents leaves ample room for continued improvement, the progress made to date evinces IRS's commitment to this aspect of the balanced measures. Moreover, the IRS has received external recognition as an excellent place to work. *BusinessWeek* recognized the IRS as one of only five non-profit / government agencies on its 2009 roster of "Best Places to Launch a Career."<sup>58</sup> Meanwhile, *DiversityInc* and *Equal Opportunity* magazines have both recognized IRS as a leader among government agencies for diversity.<sup>59</sup>

The IRS also maintains an extensive customer satisfaction program. With the institution of balanced measures, the IRS put in place transactional surveys of customer satisfaction in 11 major programs in 1998. Today, that number has grown to over 30 transactional surveys plus additional relationship surveys with small businesses, tax professionals, outreach partners, and others. Satisfaction with IRS service programs has been consistently high, exceeding 90 percent for Toll-Free and Field Assistance programs.<sup>60</sup> The IRS also participates in the American Customer Satisfaction Index (ACSI). Since 2000, ACSI has separately polled individual paper and electronic tax filers about their filing experience. From 2000 to 2009, the ACSI score for individual e-filers increased from 75 to 79 while the score for paper filers increased from 48 to 55.<sup>61</sup>

Not surprisingly, enforcement activities do not enjoy the same levels of satisfaction as service programs. Unlike service programs, these are not voluntary customers, and perceptions of satisfaction are strongly influenced by case outcomes. In SB/SE Field Examination, for instance, respondents with no tax change are four times as likely to express satisfaction as those who disagree with their adjustment.<sup>62</sup> While other factors may certainly be at play, the effect of an unwelcome outcome on overall perception of the examination experience cannot be ignored. Moreover, in contrast to the assumptions underlying the National Taxpayer Advocate's analysis, *length of process* (i.e., cycle time) has a strong influence on customer satisfaction with many IRS enforcement activities. As noted above, *time you spent on issue* is the top improvement priority for Collection customers.<sup>63</sup> *Length of process* is also identified as the top improvement priority in correspondence examinations and

<sup>56</sup> Employee satisfaction is measured as the percentage of "4" or "5" responses on a five-point scale to the question "Considering everything, how satisfied are you with your job?"

<sup>57</sup> Partnership for Public Service, *The Best Places to Work in the Federal Government*, [www.bestplacetowork.org](http://www.bestplacetowork.org).

<sup>58</sup> The Best Places to Launch a Career, *BusinessWeek* (Sept. 14, 2009), [www.businessweek.com/bestplaces](http://www.businessweek.com/bestplaces).

<sup>59</sup> The 2010 DiversityInc Top Federal Agencies for Diversity, *DiversityInc*, (Mar. 12, 2010), [www.diversityinc.com/article/7320](http://www.diversityinc.com/article/7320). 2010 Readers' Choice: Top 20 Government Employers, *Equal Opportunity*, [www.eop.com/awards-EO.php](http://www.eop.com/awards-EO.php).

<sup>60</sup> Pacific Consulting Group, *Toll-Free National Reports* and *Field Assistance National Reports*.

<sup>61</sup> The American Customer Satisfaction Index, [www.theacsi.org](http://www.theacsi.org). The ACSI score is an index, not a satisfaction percentage. The average ACSI for all individual filers has improved over time with the increased adoption of electronic filing, from 56 in 2000 to 71 in 2009. By comparison, the federal agency average was 68.7 in 2009.

<sup>62</sup> ICF Macro, *SB/SE Field Examination Customer Satisfaction Survey (April 2009 – March 2010)*.

<sup>63</sup> See Pacific Consulting Group, *SB/SE Field Collection National Report (April – June 2010)*.

large-business industry examinations.<sup>64</sup> Across an array of programs, survey results show that satisfaction decreases as cycle time increases. For example, satisfaction with SB/SE Field Examination for taxpayers whose examinations were more than twelve months long was less than half that of taxpayers whose examinations were resolved in three months or less.<sup>65</sup>

More importantly, however, the IRS uses customer and stakeholder feedback to determine how to improve the taxpayer experience to ensure that all taxpayers are treated fairly and equitably. All the surveys provide detailed response and demographic information to facilitate improvement projects. In the service context, the IRS has collaborated since 2005 with the National Taxpayer Advocate and the IRS Oversight Board to develop and implement the Taxpayer Assistance Blueprint (TAB), which drew from extensive research on customer needs and preferences to provide strategic direction to drive service improvement decisions.<sup>66</sup> On the enforcement side, we have implemented a variety of improvements, such as:

- Developing and implementing the Compliance Assurance Process, through which large-business taxpayers with a history of compliance can work jointly with the IRS to determine tax return accuracy prior to filing. This process leads to early identification of issues and facilitates faster resolution of issues than traditional examinations.
- Developing the Fast-Track Settlement process, which provides taxpayers with unagreed examination issues an opportunity to participate in a mediation process with Appeals before a 30-day letter is issued.
- Revising telephone procedures for correspondence examinations and expanding assistants' online access to all correspondence cases. These enhancements enabled toll-free assistants to provide improved service to callers by explaining the documentation required and the ramifications of not submitted the requested information.
- Producing an online video series ("Your Guide to an IRS Audit") that provides a step-by-step guide through the audit process.
- Simplifying the Form 656, *Offer in Compromise*, making it more user-friendly for taxpayers to complete, as well as pursuing other Collection and OIC improvements.

The National Taxpayer Advocate further argues that IRS measurement approaches favor funding of enforcement over service. They correctly note that "97 percent of all tax revenues the IRS collects come in without the need for expensive enforcement resources, in part because of IRS services." However, this represents only about 84 percent of the

<sup>64</sup> Pacific Consulting Group, *Compliance Center Examination SB/SE National Report (April – June 2010)* and *Large and Mid-Size Business Industry National Report (October 2008 – September 2009)*. ICF Macro, *W&I Compliance Center Examination Customer Satisfaction Survey (April – June 2010)*.

<sup>65</sup> See ICF Macro, *SB/SE Field Examination Customer Satisfaction Survey (April 2009 – March 2010)*. See also surveys of Large and Mid-Size Business (LMSB) division industry and coordinated industry examinations, and Employee Plans and Exempt Organizations examinations.

<sup>66</sup> Publication 4701, *Progress Report on the Implementation of the Taxpayer Assistance Blueprint*, <http://www.irs.gov/pub/irs-pdf/p4701.pdf> describes the guiding principles of the TAB and provides updates on the TAB service improvement portfolio. The next report to Congress is due in February 2011.

total tax liability.<sup>67</sup> Clearly, these revenues could not be collected without IRS processing and assistance operations to help taxpayers meet their obligations. Yet it would be an overstatement to brand these solely as “service revenues.” While these revenues may not be collected through any direct enforcement process, some portion of them is nevertheless an indirect effect of IRS enforcement activities. Having a credible enforcement program not only creates a disincentive to those who might otherwise attempt to evade their tax obligation, but as importantly, it reassures compliant taxpayers that the system is fair and equitable. Current enforcement actions may also have an educational result that reduces future non-compliance. These concepts are captured in the IRS’s Collecting Principles:

*Compliance – The public trust requires us to ensure that all taxpayers promptly file their returns and pay the proper amount of tax, regardless of the amount owed. The public as a whole is our customer, not just delinquent taxpayers. Our customers expect us to promote voluntary compliance by ensuring that all promptly pay their fair share. Employees should work with taxpayers to meet all their filing and paying requirements, not only the delinquency at hand. This involves identifying and addressing non-compliance to prevent future delinquencies.<sup>68</sup>*

While we agree that the vast majority of revenue is collected through voluntary compliance, it is likely that the voluntary compliance rate would drop significantly if the IRS did not have a strong enforcement presence. Without diminishing the critical value of providing quality service to help taxpayers understand and fulfill their obligations, we cannot overlook the proper role of enforcement in encouraging taxpayers to seek to comply in the first place. We agree that ensuring this balance is a delicate and sometimes difficult task, but the IRS remains committed to seeking an appropriate balance between these two goals.

While we disagree with elements of the Office of the National Taxpayer Advocate’s analysis, we agree with aspects of its four recommendations:

- **Report disaggregated satisfaction, accuracy, and completeness metrics at every level at which the IRS reports efficiency, speed, or productivity metrics.**

The IRS will continue to employ an extensive set of measures of satisfaction, accuracy, and completeness at appropriate levels of operational management. We have and will continue to review high-level documents to determine whether any changes would be beneficial.

- **Create measures for the BPR and other highly visible reports that reflect the period between the due date of the return and final resolution of any liability.**

The IRS is always looking for ways to improve its performance measurement. For example, Field Collection is already investigating ways to measure the period from

<sup>67</sup> IRS, *Tax Gap Map for Tax Year 2001* (Feb. 2007).

<sup>68</sup> IRS Policy Statement 5-2, IRM 1.2.14.1.2.

the due date of the return to final resolution for the enforcement liability inventory. However, a truly comprehensive measure would include a range of interactions (such as Examination, Collection, Appeals, and even the Taxpayer Advocate Service); it is not clear how such a measure would effectively be used for program management.

- **Create measures for the BPR and other highly visible reports that reflect the extent to which the IRS's programs completely resolved the noncompliance and prompted the taxpayer to comply in subsequent periods.**

The IRS is always looking for ways to improve its understanding of taxpayer compliance behaviors and our effectiveness. We believe the question of whether IRS programs completely resolve noncompliance and prompt future compliance is better suited to a research project than a program measure. We will consider the feasibility of future research in this area, though the effects of any particular program may be difficult to isolate from other factors that drive taxpayer behavior due to the complex nature of taxpayer behavior and the interactions of various IRS activities.

- **Initiate research to quantify the impact of various types of taxpayer service and outreach on voluntary compliance and direct tax revenue so the IRS can estimate of the ROI of such initiatives.**

The IRS has already begun a servicewide research effort (with the National Taxpayer Advocate's participation) to estimate the impact of our service and enforcement activities on the voluntary compliance of taxpayers. This effort includes developing new data on services, new ways of segmenting the population, and new ways of applying standard empirical methods to tax compliance measurement (with a particular emphasis on the role of services). However, this research is unlikely to yield specific revenue estimates but rather will improve our understanding of taxpayer needs to better support voluntary compliance.

## Taxpayer Advocate Service Comments

The National Taxpayer Advocate is pleased that the IRS agrees that the “vast majority of revenue is collected through voluntary compliance” and that these revenues could not be collected without IRS services. She also commends the IRS for improving employee engagement, for implementing the various improvement initiatives that it highlighted in its comments, and for increasing the number of customer service surveys that it conducts on a regular basis from 11 to more than 30.

Notably absent from the list of achievements highlighted in the IRS comments, however, is any significant and sustained increase in voluntary compliance, future compliance, or in the customer satisfaction reported on any of its surveys other than satisfaction with the tax return filing experience.<sup>69</sup> Nonetheless, the IRS’s comments seek to justify continuing its narrow focus on cycle time.

First, the IRS argues that it does not excessively focus on cycle time in all of its reports. It observes that some IRS reports contain a more balanced set of measures than the ones that it reports in its BPRs and Annual Performance Report. However, measures that appear in the BPR and Annual Performance Reports, which are publicly available and reviewed by IRS Commissioners, the IRS Oversight Board, and members of Congress, are more visible than other measures.<sup>70</sup> Because the measures in these reports are highly visible, they are likely to create stronger incentives than those appearing in other reports that few know about. Thus, it is important for the IRS to include a balanced set of measures in these reports.

Next, the IRS argues that even though its most highly visible reports contain a disproportionate number of measures that reflect cycle time, TAS’s analysis was “potentially misleading” because other measures are more important because of the way the IRS uses them. By analogy, a grade point average reflected in a report card may be comprised of grades from a disproportionate number of unimportant classes. The fact remains, however, that students have an incentive to get good grades in every class that will appear in their report card. They have less of an incentive to do well in classes for which they will not receive a grade (*e.g.*, classes they are auditing or taking on a pass-fail basis) or that will not appear on the report card, even if those classes are more important. IRS leaders have a similar incentive to improve the measures that appear in the BPR and similarly visible reports. If improving cycle time can raise their scores on more than one measure, they probably have

<sup>69</sup> The IRS has not measured the voluntary compliance rate since it reported the National Research Program results in 2001. The IRS asserts that the educational component of its activities, including examinations, improves future compliance, but it does not measure such future compliance, even in cases where it would be feasible to do so.

<sup>70</sup> While the Annual Performance Report is legally required (by 31 U.S.C. § 1115), IRS comments acknowledge, “[T]he 30 measures in the Annual Performance Report were selected in conjunction with Treasury to provide a high-level perspective on each IRS budget activity.” We believe this perspective should be more balanced.

a stronger incentive to put more resources into improving cycle time than other items that are supposedly more important.

Next, the IRS asserts that if it continues its focus on improving cycle time, timeliness and customer satisfaction will also somehow improve. The IRS reaches this conclusion based on surveys that suggest taxpayers want more timely service. The IRS response appears to confuse cycle time and timeliness.

Although the National Taxpayer Advocate agrees that taxpayers want the IRS to be more timely and responsive, she does not agree that continuing to focus on narrow measures of cycle time will necessarily increase timeliness or customer satisfaction. While some policies aimed at reducing cycle time may do so, others will reduce satisfaction if they leave the IRS less likely to fully resolve or address all of the issues raised by the taxpayer, or to reach the right result before closing the case.

For example, a survey of those audited by the IRS found the items with the greatest impact on customer satisfaction were: fairness of treatment by the IRS, consideration given to the information you submitted, an explanation of why adjustments were made, and the business knowledge of your auditor.<sup>71</sup> Each of these had a greater impact on satisfaction than the length of the audit.<sup>72</sup> Thus, customer satisfaction is likely to decline if IRS policies cause auditors to close cases before they can fairly consider information provided by the taxpayer, even if those policies reduce cycle time.

As the IRS response observes, the IRS's collection survey indicates that the most important driver of satisfaction among its customers is the "time you [the taxpayer – not the IRS] spent on an issue."<sup>73</sup> It would be difficult to imagine any taxpayer actually wanting to spend time dealing with an IRS collection issue. However, a taxpayer is likely to have to spend more time on a collection issue if the IRS is so focused on cycle time that it closes his or her case without completely resolving the problem or fails to offer a realistic collection alternative. As we observed in a 2004 report, the IRS's offer program once achieved remarkable cycle time improvements by simply returning 57 percent of all offer submissions to taxpayers and closing their cases without substantively reviewing the offers or resolving them in any other way.<sup>74</sup> If the taxpayer's case remains unresolved and has to be reopened later or is shifted to another IRS program (*e.g.*, from the Automated Collection System to the Collection Field Function), the IRS's narrow cycle time measures will not reflect the length of time the taxpayer had to spend to resolve the issue. As a result, IRS leaders will have little incentive to minimize it. Moreover, payment delinquencies are much easier to

<sup>71</sup> ICF Macro, *Field Examination Program, Customer Satisfaction Survey, Final Annual National Report, Closed Cases: April 2009-March 2010* 11, 32 (July 2010).

<sup>72</sup> This could be one reason that those subject to field examinations were significantly more satisfied than those subject to correspondence examinations. ICF Macro, *Field Examination Program, Customer Satisfaction Survey, Final Annual National Report, Closed Cases: April 2009-March 2010* 16 (July 2010)

<sup>73</sup> Pacific Consulting Group, *SB/SE Field Collection National Report Covering April through June 2010* L-9 (Oct. 2010).

<sup>74</sup> See National Taxpayer Advocate 2004 Annual Report to Congress 311, 314-15.

resolve if the IRS begins to work out a payment plan with the taxpayer as soon after the delinquency as possible, rather than waiting for penalties and interest to grow to unmanageable levels. For these reasons, the National Taxpayer Advocate recommended adding a taxpayer-centric measure of timeliness – the period between the return filing and final resolution of any payment delinquency.

Finally, the IRS response does not address the incentives created by the IRS's ability to estimate ROI for – and thereby obtain funding for – enforcement programs but not service programs. TAS is aware of the IRS's long-term efforts to estimate the impact of services on voluntary compliance. For the reasons discussed above, however, it is important for this (or other) research to enable the IRS to estimate an ROI for services.

## Recommendations

The National Taxpayer Advocate offers the following recommendations:

1. Report disaggregated satisfaction, accuracy, and completeness measures in the quarterly BPRs and consider adding them to enterprise-wide reports;
2. Report measures in the BPR that reflect the period between the due date of the return and final resolution of any liability (*e.g.*, through full payment, abatement, or compromise), as previously recommended by the National Taxpayer Advocate;<sup>75</sup>
3. Report measures in the BPR that reflect the extent to which the IRS's programs completely resolved the noncompliance and prompted the taxpayer to comply in subsequent periods; and
4. Ensure that the IRS is conducting the research it needs to be able to estimate the ROI of service initiatives.<sup>76</sup>

<sup>75</sup> See, *e.g.*, National Taxpayer Advocate 2006 Annual Report to Congress 62, 81.

<sup>76</sup> The National Taxpayer Advocate previously made similar recommendations. See, *e.g.*, National Taxpayer Advocate 2007 Annual Report to Congress 35, 38; National Taxpayer Advocate 2006 Annual Report to Congress 442-446.