

Occasional Paper #45

The Integration of Tax and Spending Programs

by David A. Weisbach

August 2005

Introduction

This Article provides a new theory of tax expenditures. Its argument is that the decision to implement a “nontax” program through the “tax system” has little or nothing to do with tax policy. Instead, the tax expenditure decision, which we will also call the integration decision or the decision to combine tax and spending programs, is solely a matter of institutional design. It is about assigning projects such as tax collection, education, defense, or housing to specific units of government. Different groupings of activities will perform differently, and we should use those groupings that yield the best possible performance. The problem is similar to the problem of splitting up a corporation into divisions.

Suppose, for example, that we are considering whether an education subsidy should be implemented through the tax system or through a direct expenditure. The government might use a tax exclusion, deduction, or credit implemented by the IRS or, alternatively, it might use a direct grant implemented by the Department of Education.

The two leading theories that purport to address this question of how to allocate government largesse both focus on tax policy. The most widely accepted theory, the comprehensive tax base theory, argues that a broad tax base distorts economic decisionmaking less than a narrow base and is also much simpler to administer. To ensure that the tax base is as broad as possible, this theory suggests that spending and regulatory programs should not be implemented through the tax system; instead, they should be assigned to other agencies or departments. Integrating a spending program into the tax system, for example through a deduction, exclusion, or credit for some special activity, narrows the tax base and makes the tax system more complex. The standard or default response according to a comprehensive tax base theory, therefore, is that government spending programs, such as the education subsidy in our example above, should not be implemented through the tax system.

The other leading theory, the theory of tax expenditures, focuses partially on institutional design, but it, too, is ultimately a theory of taxation. The key insight of this theory is to recognize the functional equivalence of putting a program in the tax system or somewhere else. While this insight relates to institutional design, the theory ultimately falls back on tax policy for recommendations. For example, the distributional effects of a policy are said to depend on whether it is correctly considered part of the tax base. The tax expenditure theory, just like the comprehensive tax base theory, would conclude that the education subsidy should not be implemented through the tax system.

In contrast to these theories, which focus on taxation, our theory focuses on institutional design—the question of how the government chooses to compartmentalize its functions. It is entirely irrelevant whether some piece of government policy complies with independent tax norms. If the underlying policy is held constant, there are no effects of putting a program into or taking a program out of the tax system even if doing so hurts or enhances traditional notions of tax policy. Welfare is the same regardless of whether the program is formally part of the tax system or is located somewhere else in the government. If we mistakenly look only at the tax system instead of overall government policy, we will draw the wrong conclusions. Putting a program into the tax system makes the tax system look more complicated, but there is unseen simplification elsewhere. The tax system will seem less efficient, but the efficiency of government policy is unchanged.

The institutional design question is about dividing the government into units that will provide the best possible set of public policies and government services. Different groupings of government services will perform differently. Consider, for example, a proposal to have the IRS run the country's defense system, replacing the Department of Defense. The proposal is not as silly as it sounds. It would not mean that bespectacled revenue agents would be parachuting into the Hindu Kush wearing night goggles, camouflage, and pocket protectors. Instead, an intelligent Commissioner of Internal Revenue would allow his employees to specialize. Revenue agents would specialize in reading financial statements and soldiers would specialize in fighting. Policies under such a proposal might very well continue much as they do today. The reason why the proposal is not a good idea is that there are no benefits to coordinating these two functions of tax administration and defense. Moreover, there would be additional costs because at the very top level, where functions must be combined, administrators would be unable to specialize in these relatively distinct functions.

Consider instead a proposal to implement all federal welfare-type programs through the IRS. Proposals of this sort have been made frequently, often under the rubric of a negative income tax. This may make sense, if there are benefits to putting welfare and tax into the same organizational unit: Both programs rely on income or wealth measurement, both need large-scale information and financial processing, and both have substantive policies, such as the policy of redistribution, that overlap significantly. The two programs thus might benefit significantly from coordination by a single agency.

The key variables from this perspective have nothing to do with tax policy. Instead, they have to do with the benefits of coordination between and specialization within various types of activities performed by the government. The Department of Defense needs highly specialized operatives, and thus benefits little from coordination with the revenue collection function. Welfare programs, on the other hand, may gain much from coordination with tax collection, and there may be low costs to losing the utility of separate units that can specialize in each function. The question is one of tradeoffs between the benefits of specialization on the one hand and the benefits of coordination on the other.

This intuition strongly contrasts with the usual tax arguments. For example, the Flat Tax is an attempt to provide a comprehensive consumption tax base. All the extraneous, nontax elements of current tax law would be removed. The Flat Tax is said to be very simple, and it may be if one looks only at the tax system. But limiting the tax system to this one measurement would force other

government programs to take up the slack: Programs of all sorts that are now embedded in the tax system will have to be implemented by other government agencies. Viewing the Flat Tax (or any comprehensive tax base) as simple requires ignoring the rest of government, relegating the complexity and mess of government spending and regulation to somebody else's backyard. There is no reason to believe, however, that the tax collection function should necessarily be separated from other functions of government—and there may be good reasons to believe that it should not be. The same is true for virtually all proposed fundamental tax reforms.

This Article expands on these intuitions. Part I frames the question as one of institutional design rather than tax policy. Part II discusses the comprehensive tax base and tax expenditures literatures. We argue that neither literature provides a convincing answer to the question of how government spending programs should be organized. The comprehensive tax base argument is the more prevalent of the two, but it ignores the basic problem of organization: It takes a completely tax-centric view of government and, therefore, leads to faulty conclusions. The tax expenditures literature addresses the problem—at least to some extent—as one of institutional design, but its analysis is neither complete nor convincing.

Part III approaches the problem from an institutional design standpoint. The study of organizations is old and deep, extending into sociology, economics, political science, and even anthropology and psychology. Covering even a small portion of this literature is well beyond the scope of one paper, but the problem can be divided into three more manageable pieces. First, we can view organizations as devices for coordinating specialized functions and, in particular, for separating production processes into tasks or divisions in the most efficient manner possible. Second, we can view organizations as a way to solve agency problems. Third, we can view the design of public organizations as a method of resolving public choice problems. The volume of literature in each of these areas is very large, and the more informal literature generally mixes these areas together. The portions that relate directly to the problem of divisions, however, are reasonably manageable and in some cases quite sparse. In this Article we address only the first of these three questions—the specialized production part of the puzzle. We intend to address the agency and public choice elements of organizational theory in subsequent work.

The key intuition on specialization of production is that there are benefits to both specialization in particular activities and coordination between activities. Putting a set of activities into a governmental agency promotes specialization within that set of activities and coordination among the activities. But at the same time, doing so makes it more difficult to coordinate between the activities in that agency and the activities of other agencies. Part III develops this intuition by surveying the literature in the area of organizational structure.

To demonstrate the worth of these intuitions, Part IV works through several examples, evaluating government programs from the standpoint of institutional design. In particular, it examines two federal programs—the food stamp program and the Earned Income Tax Credit (EITC)—and considers whether either should be implemented as part of the tax system. There are good reasons to think both should be: They are both transfer programs based on income, and the IRS may be the agency that is best able to perform this income-based redistribution. Nevertheless, we conclude that the food stamp program may not function as well if it is made part of the tax system, because the tax system cannot respond in a sufficiently timely fashion to the needs served by the food stamp

program. The EITC, by contrast, works reasonably well within the tax system. The difference between the two programs is the degree to which their functions complement those performed by the tax system. Part V concludes.

I. Framing the Question

The central question of this Article is how to determine the best way to implement a government program, given that such a program is going to be implemented. As Stanley Surrey noted in his tax expenditures analysis, virtually any program can be implemented in at least two ways. It can be implemented through a direct spending program or through a tax program. The question is how to make this choice.

For example, suppose that the government wants to provide an education incentive. The incentive might, for example, take the form of a voucher program, providing money for students to use in choosing high schools. Alternatively, the incentive might try to reduce the cost of saving or borrowing money to pay for college, as do the recently enacted college savings plans and income-contingent student loans. In either case, the grant might be based on attendance at an appropriate institution, income, citizenship, race, lack of criminal convictions, grades, or a variety of other attributes.

Such a program can be implemented through an expenditure program that distributes money to individuals meeting the criteria. The department implementing the program, say the Department of Education, would have to create an application process, a certification or audit process (both for students and schools), a process for handing out money, and, if appropriate, a process for collecting payments. Setting up such a program would be complex and would take significant resources. Alternatively, a similar program could be implemented through the tax system by allowing individuals to subtract or add the same amounts to their taxes (or if their taxes are not sufficient, by requiring the Department of the Treasury to write a check to the individual based on a claim made on his tax form). Similar application, certification, and auditing requirements could be imposed. The additional requirements imposed on the tax system would also be significant, making the tax system much more complex. As a final alternative, the program could be split between the tax system and another agency, with each agency providing services related to its expertise and some coordination between the operations.

Similarly, suppose the government wants to provide welfare to a subset of the poor. For example, the government may want to provide welfare to those poor who work a certain amount. Such a welfare program can be implemented through the tax system. The EITC does exactly this: It provides a tax credit for individuals if their labor earnings are within a certain range, subject to a variety of other criteria (for example, different credits are granted to the married, the unmarried, those with and without children, and those with earnings from sources other than labor). A similar program could also be implemented outside the tax system, through a direct grant of aid based on similar criteria.

The question is how we should decide which is the better method of implementing these programs in each of their circumstances. The key is to assume that a program of some sort will be implemented, so that it is not an option to say that these are bad ideas and we should do nothing.

They may very well be bad ideas, and actual programs implemented by the government may be even worse, but unless one is going to admit no role for the government other than the most minimal functions, these sorts of programs and problems will arise. The government will, sometimes for the better and sometimes for the worse, subsidize, penalize, or regulate various activities, and we must decide how this should be done.

We need not belabor the importance of the question. As Surrey—and others before and after—have established, there are a vast number of programs implemented through the tax system. Both the executive and Congress make annual estimates of the size of these expenditures, and the amounts are large. For example, the Bush Administration's 2004 budget estimates that in 2003, the exclusion of employer-provided contributions for medical insurance reduced tax revenues by \$108.5 billion, and the deduction for mortgage interest on owner-occupied homes cost \$65.5 billion. The same budget lists fourteen different education programs in the tax code. The EITC is one of the largest welfare programs, having grown faster than any other program over most of the last decade. Whether programs such as these should be implemented through the tax system is thus an important question, and one that has commanded significant scholarly attention.

There are four important limitations to our framing of the question. First, we generally will discuss only comparisons between direct grants and tax subsidies, although a similar analysis could be applied to regulatory programs. Regulatory programs can be implemented through tax or expenditure programs: For example, although we could directly regulate executive salaries, we instead impose tax penalties for executive salaries that do not conform to certain requirements. Similarly, tax and expenditure programs can be implemented through regulation: The minimum wage, for example, can be viewed as a combination of a tax and a spending program—taxing employers and providing grants to employees—that has been implemented through regulations. We are not even sure that there is a clear dividing line between spending, regulation, and taxation, and we do not wish to imply that the analysis of government programs must be limited by any such taxonomy. Nevertheless, our examples tend to focus on the comparison between direct grants and taxes, and there may be differences in the context of regulatory programs that we gloss over.

Second, there are both bureaucratic and legislative components to the implementation of any program that typically (but not always) move in tandem. Putting a program into the tax system generally (but not always) involves delegating the program to the IRS and also at least partially delegating legislative jurisdiction to the tax-writing committees. Putting the program into another agency generally involves delegating legislative jurisdiction to the appropriate oversight committee for that agency.

The interaction between the committee structure, the legislative floor, the agency, and the President can be very complex. Problems of legislative organization raise unique questions because legislatures cannot choose their members and have a low tolerance for hierarchies. The internal design of a legislature might, therefore, differ from the stand-alone optimal design of a bureaucracy. The interaction of legislative design with oversight requirements may affect bureaucratic design.

To simplify the problem, our focus will generally be on agencies rather than other political actors. Principal-agent problems between the legislature and agencies, between the Executive and

agencies, and within the legislature are likely to be a central piece of the integration issue, but in this Article we focus only on how to organize specialized production across potential implementing agencies, ignoring these principal-agent problems for now.

The third simplification is that we will ignore agency capture, interest group activity, and other problems of public choice. Agency capture is in a sense just a variant on the principal-agent problem, where the agent's preferences coincide with the preferences of the regulated industry and do not align with the preferences of the principals. The difference is that agency capture arises because of problems of collective action (such as monitoring of the agency by large groups of individuals). The same is true of many of the public choice problems that arise with respect to agencies. These issues have important consequences for the design of agencies; for example, scholars have considered the use of tax expenditures as a way to limit capture of congressional committees. While agency capture and other public choice considerations are likely to be important in considering whether to put programs in the tax system, we put the issue to one side.

Fourth, we generally take the limits and boundaries of current institutions as given. We assume, for example, that both the tax system and the education system exist as they are, and that the question is how to implement an educational program, rather than how to restructure one or both systems to make them run more efficiently. The framework we set forth could apply more generally to wholesale reorganizations of the government and its agencies, but in our examples, and to some extent in our thinking, we have confined ourselves to a smaller goal—namely, to determine how to allocate a program across existing institutions, treating the merits and flaws of those institutions as largely fixed.

To summarize, we frame our question as a choice among implementation methods for a program that is going to be adopted. The choice involves picking which government agency or agencies should implement the program, with the primary focus on whether it should be the tax system or some other agency. We focus on the organization of production within the government bureaucracy and the effects of coordination and specialization—leaving aside questions about the legislature, principal-agent problems, and agency capture.

II. Do the Comprehensive Tax Base and Tax Expenditures Literatures Answer the Question?

A. The Comprehensive Tax Base

The goal of comprehensive tax base (CTB) advocates is to provide a broad tax base that has few or no exceptions, preferences, or loopholes. CTBs generally come in two flavors: income CTBs and consumption CTBs. An income CTB attempts to tax some comprehensive measure of income, the details of which depend on each advocate's taste for purity over administrative complexity (or any other sufficiently important consideration). A consumption CTB attempts to tax a comprehensive measure of consumption.

The arguments favoring either an income CTB or a consumption CTB are based on both efficiency and fairness. The efficiency argument is that a broad tax base is more neutral between activities than is a narrow base. For example, the current income tax does not tax owner-occupied homes but does tax corporate capital. By providing a relative disincentive to invest in corporate capital over

owner-occupied homes, these differential tax rates change where individuals allocate capital and, therefore, distort markets. A CTB would be neutral between investments in owner-occupied homes and corporations, and would therefore lead to a more efficient allocation of capital. Moreover, with fewer “leakages” tax rates would be lower, which would mean that inefficiencies would be further reduced. The fairness argument is similar: Individuals would be taxed the same regardless of their preferences for engaging in different sorts of activities.

The CTB agenda, if followed, almost surely would lead to substantial improvement in government policy. Many, if not most, of the exemptions to the tax base are unwarranted, and broadening the tax base will often lead to a more efficient, fairer, and simpler set of policies. We have no quarrel with much of the agenda of CTB advocates and applaud much of their work.

The CTB literature, however, has nothing to say about the question posed here. If we are going to have a program and the only choice is how it is going to be implemented, the efficiency and fairness arguments in the CTB literature completely fail. If we hold the content of the policy fixed, the efficiency and fairness arguments in the CTB literature completely fail. If we hold the content of the policy fixed, the efficiency implications are the same regardless of whether the tax agency or some other agency implements it. Similarly, the fairness of the policy is the same regardless of whether the tax agency or another agency administers it. If one considers a program that is going to be implemented one way or another, the conclusions of the CTB literature may be completely wrong. There is no a priori reason to believe that a broader tax base is better in such a situation.

It is possible that CTB advocates believe that government should never subsidize, penalize, or regulate activities. Broadening the tax base, then, would not cause similar programs to be implemented elsewhere. Instead, it would cause them to be eliminated. In many cases this may be desirable and a good description of the effect of broadening the tax base. For example, many of the base-broadening provisions of the 1986 tax reform did not result in the creation of substitute programs in other agencies. But this view is unlikely to be either desirable or a good prediction about government policy in all cases. If we are going to have a program, the CTB literature simply has nothing to say about where it should be administered.

As far as we can tell, the only way one can make the arguments made by CTB advocates is to treat the tax system as separate from the rest of the government. Under this view, ensuring the efficiency and fairness of the tax system, taken alone, should be the goal of tax reformers. This produces the odd result that removing a program from the tax system and replacing it with an identical program implemented by another agency produces an efficiency and fairness gain, notwithstanding that no behavior has changed and no policy has been altered. Kyle Logue has called this view “tax-exceptionalism.” Another name might be NIMBY, or Not In My Back Yard. CTB advocates want their backyard to be clean but don’t seem to care about where else the trash might be put.

Some CTB advocates make a political (as opposed to an efficiency or fairness) argument against integrating spending programs into the tax system. The idea is that one tax loophole leads to another. As Walter Blum puts it,

[please inset the following paragraph and delete this instruction, thx]

There is nothing about the combination of rate reduction and base broadening which dictates that all preferential provisions be eliminated, but there are potent reasons for leaning over backwards before allowing any of them to remain in the law... . The existence of any one special dispensation makes it easier to argue on behalf of others... . [A] Spartan attitude toward defending the integrity of the base will aid in creating the impression that the reform plan is intended to improve the system as a whole, with the chips falling as they may, and is not calculated to benefit certain identifiable groups possessing political strength.

This argument, however, is unpersuasive. Shifting programs from the tax system to other parts of the government does not change the amount of government largesse. If handing out goodies to one group makes it difficult to say “no” to another group, then putting a program in another agency does little to make it easier.

One possible reason that it may be more difficult to limit largesse in the tax system may be that tax programs are less visible than direct spending. This argument is at least on the right track, in the sense that it compares different ways of implementing a program. The visibility argument, however, is not convincing. Between the tax expenditure budget and the wide variety of articles and books discussing tax breaks, there is no reason to believe that individuals are not as well-informed about tax breaks as direct subsidies. In many cases, it is hard to believe that tax expenditures are less visible than other government programs. For example, there is little reason to believe that the home mortgage interest deduction is less visible than, say, the implicit guarantee the government provides to Fannie Mae and Freddie Mac to lower mortgage costs. It is similarly unclear why tax deductions for brownfields cleanup would be less visible than any implicit subsidies one might find in other environmental rules.

Perhaps endowment effects make expenditures through the tax system less visible than direct expenditures. People may perceive a reduction in taxes for engaging in a specified activity differently from an identical direct grant: They may perceive a tax subsidy as merely letting them keep their money, even while they perceive an identical program that taxes them and gives the money back through programs or services to be a subsidy. Attempts to publicize the extent of tax subsidies through budgets or books will not be able to overcome this flaw in our reasoning ability.

We are not sure of the extent to which such a flaw exists or is a problem. There does not seem to be a framing effect among those who study or work with the tax system; lobbyists and their clients fully understand the benefits of tax credits or deductions. If there is a framing effect, it can be overcome. The argument is really that the hoi polloi are too stupid to understand the equivalence between a tax benefit and direct spending. But if the argument relies on stupidity, it is hard to see how the same individuals will understand or even know about the vast number of direct spending programs, many of which have very subtle and indirect effects.

For example, it is hard to believe that individuals understand the subsidies for driving created by federal highway spending any better than they understand the subsidies for driving created by tax benefits for oil companies. Both are very complex programs that only indirectly affect the consumer. Neither shows up in a form or application that individuals see. Most individuals probably never think about the huge subsidies given to their automobiles. To the extent that they do, it is hard to believe that they understand the direct expenditure better than the tax expenditure.

Psychological problems may prevent individuals from properly processing information, but this does not mean that such cognitive biases are dominant, or even important, in this context. Moreover, it is not clear in which direction these biases operate.

To the extent we believe the visibility argument, it may actually lead to a legislative preference for direct spending over tax programs rather than the other way around. For example, if a congressman can fight for either tax reductions or direct spending, and if constituents perceive direct spending dollars to be greater than equivalent amounts of tax reductions, then the congressman may prefer direct spending. The congressman gets more constituent bang for the same budgetary buck. In addition, direct spending may be socially preferable for rent-seeking programs, since it will take fewer direct spending dollars to satisfy the rent seeker who is subject to framing. Alternatively, lack of visibility may be a good thing rather than a bad thing. If the program is desirable but also one that individuals tend to resist (like eating your vegetables), putting it into the tax system could reduce opposition by making it invisible due to framing effects. There is no general theory of psychology and politics implying that the comprehensive tax base is either a good or a bad idea.

To summarize, if programs are going to exist, the CTB literature has nothing to say about where or how they should be implemented. To be fair to CTB advocates, many of the base-broadening (and base-shrinking) proposals make sense, and we do not mean to claim that the literature is not valuable. But at the same time, CTB theory focuses exclusively on the tax system, essentially assuming the answer to the integration question.

B. Tax Expenditures

The tax expenditures literature, although related to the CTB literature, focuses directly on the integration question. It offers many potential insights into that question, but it is also seriously flawed. We break our discussion of tax expenditures into three parts. First, we briefly review the basic idea of tax expenditures. We then review the chief criticism of the idea, the problem of the definition of tax expenditures. Finally, we discuss the merits of the substantive arguments made by the literature, concluding that one of the core intuitions has merit but that most of the details are unconvincing.

The basic idea behind tax expenditures is that any government program can be implemented through a direct expenditure or through the tax system. Any time the government uses the tax system as opposed to a direct spending program, the government has created what is labeled a “tax expenditure.” The tax expenditures literature, primarily through the writings of Stanley Surrey, did three things with this insight. First, it tried to define the notion of tax expenditures. The core definition is that any deviation from a “normative income tax” is a tax expenditure. Thus, any deduction, exclusion, or credit that would not be allowed under some definition of income is equivalent to a direct expenditure. Second, it argued that the budget rules as then in effect distorted the process toward tax expenditures, particularly by hiding information about the costs of tax expenditures and by having more lenient procedures for enactment. The goal of the tax expenditures literature in this regard was to create a budgetary accounting for tax expenditures that mirrored that of direct expenditures, thereby reducing the budgetary incentives to use tax expenditures. Third, the tax expenditures literature discussed the merits of tax expenditures, generally concluding that they are an inferior method of implementing policy.

The definition of tax expenditures has been frequently debated in the literature. A tax expenditure is said to be any deviation from a “normative” tax base. The normative tax base is defined as an amended version of the comprehensive income tax base, but the particular details vary by individual tax expert. In the theory of tax expenditures, much rides on such differences in definition: If a particular deduction or credit falls within the bounds of the normative tax base, none of the consequences of being a tax expenditure applies, while if it falls outside of the normative tax base, all of the consequences apply.

Bittker is the most prominent critic of this approach. He argues that a comprehensive definition of income would include many items not contained in the “normative tax base” of tax expenditure proponents. For example the tax expenditure supporters do not include the benefit of the cash method of accounting, the realization requirement, and imputed income from assets and housework in their normative tax base. Moreover, Bittker argues that even among supporters of a tax expenditure theory, there is no widely accepted definition of the proper tax base. For example, there is no broad consensus regarding the acceptable degree of progressivity or of the scope of a family.

Bittker is not merely accusing tax expenditure theorists of inconsistency in defining tax expenditures. Instead, he is accusing them of making implicit policy judgments. If all of these exceptions from the tax expenditures list are based on policy judgments, then so are the items on the list. Policy judgments, however, do not come from definitions. Instead, Bittker concludes that, short of a coherent conceptual model, each exemption from the tax base must be examined and evaluated on its own merits.

Stated in the language we are using here, it is hard to see how the organization of a bureaucracy should depend on a definition of income. For example, if we are going to subsidize medical expenses, whether it is desirable to do so through the tax system should not depend on whether a medical expense deduction meets the definition of income. Debates about the matter seem completely beside the point. Similarly, Surrey’s upside-down subsidy argument, discussed below, only applies to items not meeting the definition of income. It is hard to imagine that the distributional effects of a provision depend on meeting a definition.

Another way to see the problem with the definitional approach is to consider the integration question with respect to nontax agencies, as suggested by Professors Douglas Kahn and Jeffrey Lehman. Suppose, for example, we are trying to determine whether a conservation program should be put into the Department of the Interior, the Environmental Protection Agency, the Army Corps of Engineers, or somewhere else in the government. The answer depends on institutional factors, not on definitions. The same should be true for deciding whether a spending program should be put into the IRS.

The most heated offshoot of this debate about definitions is the debate about the scope of the tax expenditure budget. At the time Surrey was writing, tax programs and direct spending programs were treated very differently by the budget system (and they still are). The budget reported the amounts spent on direct expenditures but did not report amounts spent through virtually identical programs in the tax system. The differences in the budget rules were liable to distort outcomes. To

remedy this, Surrey called for a “full accounting” of tax expenditures, and the result was the tax expenditure budget, which lists the costs of various items in the tax law.

Some sort of definition of tax expenditures is necessary to have a tax expenditure budget: We must have some method of measuring tax expenditures if they are to be reported in the full accounting. As noted above and by numerous commentators, however, there is no a priori definition of the tax system. There is no such thing as a normative tax base.

The simultaneous need for a definition and the lack of grounding for any particular definition make the tax expenditure budget problematic. For example, we must decide whether accelerated depreciation is a tax expenditure. It provides faster cost recovery than economic depreciation but slower cost recovery than expensing. It can alternatively be viewed as a tax expenditure or a tax penalty depending on whether one’s baseline is an income tax or a consumption tax. Critics claim that such problems with definitions are fatal to the exercise of constructing a tax expenditure budget.

Perhaps the reason for the heated debate is the normative consequences associated with labeling a tax provision as a tax expenditure. Surrey’s arguments, as well as those of the comprehensive tax base literature, generally condemn tax expenditures, so such labeling becomes extremely important. Being put on the tax expenditures list indicates that a provision is a subsidy or government largesse, while staying off that list indicates that a provision has the patina of good tax policy.

If we reject the normative consequence of the label, however, the problem becomes much simpler. The problem is one of determining what information would be useful. For example, it would be useful to know both how much tax revenue would go up if accelerated depreciation were to be replaced with straight-line or economic depreciation and how much it would cost to replace it with expensing. One does not have to answer the question about which way of looking at the issue is right.

The decision about what information to release will be difficult and problematic. It will inevitably have normative and political components. But it would be a significantly lesser problem than deciding what to include in the one true tax expenditures list. For example, information could be presented in a variety of nonevaluative ways under an “information usefulness” rationale, whereas there is only one correct (and strongly evaluative) method to present the same information under a traditional tax expenditures rationale. Similarly, as has been suggested by others, information would only need to be provided for those elements of the tax system that could conceivably be replaced with a direct expenditure program, because only these programs can be distorted through budgetary nonneutrality. Surrey’s argument that the budget process could skew outcomes was correct; we should not let unrelated problems with his arguments get in the way of this truly valuable contribution.

Much of the debate about tax expenditures has focused on these definitional problems. Although Surrey apparently believed and cared about the definitional parts of his argument, this focus on definitions distracts from the underlying substantive arguments surrounding tax expenditures. Once definitions are put aside, the tax expenditures question really is the integration question. We believe

that many of the substantive arguments made in the tax expenditures literature about integration are unconvincing, but that there is a core, unstated intuition that remains valuable.

Surrey's substantive argument is that tax expenditures are an inferior method of implementing policy. Instead, he claims, the government should use direct spending programs. The reason is that tax expenditures tend to have a variety of features that lead to poor implementation. For example, Surrey argues that tax expenditures create so-called upside-down subsidies. Upside-down subsidies are created because the value of tax deductions increases with the marginal tax rate, so that wealthy individuals with high marginal tax rates will receive more for a given deduction than individuals with lower incomes and lower marginal tax rates. If one views tax expenditures as equivalent to the government handing out money, wealthy individuals get bigger handouts than the poor.

Tax expenditures also possess other traits that tend to make them open to inefficiency or abuse. First, they place no limits on the amount of tax benefits a taxpayer may receive, and hence are not capped. Second, because they are part of the Internal Revenue Code, tax expenditures do not require an annual appropriation (as agency programs do); rather, they are like direct expenditures that are automatically appropriated absent some contrary congressional action. Third, tax expenditures tend to have relatively loose eligibility requirements, in the sense that individuals self-declare their eligibility and are challenged only if they happen to be audited. And finally, as Surrey argues, tax expenditures in general create more complexity than direct expenditures.

The immediate response to such arguments is that we can overcome these flaws by designing tax expenditures to be implemented in the same manner as direct expenditures. Indeed, Surrey begins his argument by assuming that tax expenditures could be applied in a way that is basically identical to direct expenditures. If direct and tax expenditures have the same content, however, these criticisms do not apply. For example, if a tax expenditure has the same content as a direct spending program, it will not have the upside-down subsidy effect, it will not be open-ended, its eligibility criteria will be the same as those of a direct expenditure, and it will not be more complex than the direct spending program. Moreover, many new tax expenditures are designed to be more similar to direct spending programs than prior tax expenditures were. Congress now tends to use credits rather than deductions and has, where it has thought it appropriate, limited the size of the expenditure or the eligibility for the expenditure.

It is curious that someone as sophisticated as Surrey would make such an obvious mistake. While Surrey states that he wants to compare identical tax and direct expenditures, if he truly did so, his conclusion would have to be that the two programs are identical. Instead, Surrey concludes that direct expenditures are superior to tax expenditures. He does so by comparing different programs, notwithstanding his assertion that the programs can be made identical. Yet he does not justify why he compares different programs.

Perhaps Surrey compares different programs merely because that is what he observed. But we believe that there is a key, unstated intuition driving him toward this approach: Institutions matter. If institutions matter, policies will be different when implemented by different institutions. Thus, tax expenditures and direct expenditures will tend to have different features and should not be compared as if they were identical.

For example, the use of exclusions or deductions to implement tax expenditures may create an upside-down subsidy effect, but doing so also has an important offsetting feature: An exclusion is an incredibly simple method of implementing policy and a deduction is only a little bit more complicated. Similarly, self-declared eligibility is much simpler than other methods of implementing policy.

While these features could be made part of a direct spending program, they are a key benefit of integration into the tax system because they take advantage of the existing infrastructure of tax collection. That is, integration allows for economies of scope in policy implementation with the resulting savings in administering and complying with the system. We would not necessarily want to design tax expenditures to be the same as direct expenditures. If we did, we would defeat the purpose behind putting a program into the tax system. We would lose the benefits of integration.

For example, if we want to subsidize the development of human capital, we can create direct spending programs that might subsidize training or education. An alternative is to defer taxation of the returns to education. Exclusion and deferral have many flaws, but a key—perhaps decisive—advantage is that they are incredibly simple. The current system is transparent to individuals. Integration of the education subsidy with the tax system achieves this transparency better than a direct spending education program with identical substance.

One way to frame the choice between tax and direct spending is to begin with a set of broad policy objectives and allow complete freedom in designing a program to meet them. The objectives can be met in a variety of institutional settings, and we should compare the best possible program that can be implemented in each of these different settings. Integration, for example, allows coordination of programs and the use of a common infrastructure for administering programs. But in integrating two programs, we might give up making each of them as accurate as possible with a specialized agency. We choose the method of implementation that offers the best results.

Surrey, then, had the key idea right even if it was not explained: We should not compare identical programs when making the integration decision. We should compare programs that are best designed for each institutional structure and choose the best from among these.

Consider for example, Anne Alstott's discussion of the integration of welfare programs with the tax system. She focuses on whether the EITC, which can be viewed as a substitute for more traditional welfare programs, is a good idea, but her arguments have broader application. Integrating welfare and tax systems may greatly simplify government policy by utilizing the existing institutions of the federal income tax. But the benefit of a stand-alone welfare program would be that it could test eligibility and provide benefits over short periods, in order to account for the fact that individuals' welfare needs can vary dramatically over short periods of time. If one uses the institutional structure of the tax system to implement a welfare program, one must almost inevitably use an annual accounting period. Therefore, one of the tradeoffs of integration is that the program is less well-tailored to its needs—that is, it will be less accurate. The tradeoff is between the simplicity benefits of integration and the accuracy benefits of separation.

Note that this turns the usual complexity/simplicity argument on its head. The usual argument is that putting programs into the tax system increases complexity. This argument is correct if one

looks only at the tax system. But if one considers government policy as a whole, integration with the tax system may often be a choice for simplicity. Integration is a choice to take advantage of the infrastructure of the tax system at the cost of less accuracy in program design than would be achieved through a separate agency.

Surrey's arguments do not hold up well under this type of analysis because he does not consider the benefits of the design features that he observes. For example, the open-endedness and eligibility declaration criteria that he criticizes greatly simplify the system. He argues that we do not find these features in direct expenditure programs, so they must be undesirable. But whether they are found in direct expenditure programs is entirely beside the point. The decision to put a program into the tax system can be seen as a decision that the accuracy costs of these features are outweighed by the simplicity benefits of integration for those particular programs. We should expect to see different features in tax expenditures and direct expenditures. In fact, we should expect to see different features in tax expenditures and direct expenditures for precisely the reason that Surrey provides in comparing different programs—because design features will and should vary with the institution implementing the program.

To summarize, the tax expenditures literature focuses on the integration question and does not take a NIMBY-type attitude, unlike the CTB literature. It gets confounded in unnecessary definitional debates, but it provides a key insight: When comparing methods of implementing policy, we should not compare identical programs, but instead should compare how a policy is likely to be implemented in any given institutional structure. The contribution of this framework is significant, but the arguments of the tax expenditures literature are unconvincing precisely because they fail to take full advantage of this framework. Tax expenditure theory fails to account for the inherent benefits of integration, and instead focuses on a tax-centric consideration of complexity. In so doing, tax expenditure theorists blind themselves to the differences that matter most between tax and direct expenditures: the simplification that the tax system provides on the one hand, and the tailoring and accuracy that direct spending programs provide on the other.

III. The Integration Question as Organizational Theory: Coordination and Specialization in Production

We view the integration question as one of how best to organize the government rather than a question of tax policy. This means that the relevant place to look is in the organizational literature rather than in the tax policy literature. The literature on organizations is vast, going back to Taylor's scientific management, Weber's studies of bureaucracy, Coase's theory of the firm, Simon's theory of administrative organizations, Chandler and Williamson's M and U theories, Marschak and Radner's theory of teams, and Niskanen's models of self-serving bureaucrats. It continues today in modern information processing models, agency models, and positive political science models.

We address here only a relatively narrow question of organizational theory. The question we must ask is how we should optimally split up an organization into divisions. If one thinks of a hierarchy as an upside-down tree, the question is the placement of vertical lines that split it into divisions or agencies. Much of the literature is on the number of layers or horizontal lines, which is related but not directly on point. Other portions of the literature on the nature of leadership within

organizations are also not relevant. Theories of the firm are about boundaries, but the boundaries are usually between the market and the firm, not within the firm. It turns out that only a very small portion of the literature focuses on the issue of optimal divisions within an organization.

We can break the applicable literature into three parts. The first part includes theories that assume away any divergence of preferences among individual agents. Following Marschak and Radner, we generally refer to this line of literature as team theory. Hierarchy arises in team theory because of limitations on information processing. Hierarchies split up the decisionmaking or information processing tasks into subunits and allow coordination of these subunits through higher tiers in the hierarchy. The shape of the hierarchy usually depends on the gains from specialization and the costs of coordination. Most of the team theory literature, however, focuses on the “horizontal” dimensions of hierarchy, such as the number of layers in a hierarchy or the span of control for a manager. Our problem is how to divide the organization into divisions rather than how many layers of management it should have. Nevertheless, the intuitions from the literature about the tradeoff between specialization and coordination may be applicable.

The second and largest part of the literature focuses on incentives. This literature assumes agents have some specialized knowledge that makes delegation attractive and, in addition, that agents have preferences that diverge from those of their principals. The goal is to set up a hierarchy that gets the benefit of organized production (e.g., agents’ expertise) while minimizing shirking. We can think of most tax expenditures as split delegations, where part of a policy is delegated to a specialized agency and part to the IRS. For example, much education policy is administered by the Department of Education, but some is administered by the IRS. Similarly, welfare is split between specialized welfare agencies and the IRS. The integration question thus can be seen as a question of when the use of multiple agents is desirable.

The last portion of the literature focuses on collective action or public choice problems. Much of this work is closely related to the agency literature in that it generally assumes that a politician or bureaucrat cannot be fully monitored by the public, but the focus of this branch of inquiry is on the problem of decisionmaking or monitoring when there are a large number of diverse principals rather than one or a few principals.

In this Article we focus only on team theory as applied to the problem of divisions. The question is how we can organize a bureaucracy to best facilitate specialized production when there is no divergence of preferences between agencies, the individuals who make up the agencies, and the legislature or principal. This Part begins by addressing the problem in a general context and then by turning to how the analysis applies to bureaucracies and tax expenditures. The next Part applies the analysis to two examples—Food Stamps and the EITC.

A. The General Problem of Divisions

There is a tradeoff in deciding whether to assign a function to a separate division. Separating a function from the rest of the firm in this manner allows specialization and coordination of activities within the division, but increases coordination costs between that division and other activities of the firm. Think of a hotel company that separates the operation of its luxury hotels from its value hotels. The separation allows each division to specialize in providing a particular type of service

and to coordinate those activities cheaply. But the separation means that coordination between the divisions will be more difficult. If coordination among activities is particularly important—say, because adopting common standards or operating procedures will save costs—the benefit of specialization and coordination within an activity may not be worthwhile. Conversely, the more valuable specialization and coordination are within a group of activities, and the less important coordination is among different groups, the more likely it will be a good idea to separate those groups.

Much has been written on these questions, going back at least to Adam Smith. Specialization adds value because it allows an individual or organization to perform the same activity more rapidly, more accurately, or better in some other dimension. Smith uses the example of a pin factory. A single individual can make very few pins in a given time period. A group of individuals, each specializing in a single aspect of making pins, can vastly increase output. The reason why is that specialization in particular elements of pinmaking allows individuals to perform their particular tasks more efficiently.

The key question is what limits specialization. Why not have a separate division for each individual function performed by the firm? The answer, as illustrated by Becker and Murphy, is that specialization is limited by the costs of coordination. Too much specialization means that coordination of the specialized activities becomes difficult. For example, pediatricians generally do not specialize in particular childhood diseases. While they would learn more about a disease through specialization, the additional knowledge would require greater expenses in coordinating their care with other pediatricians. The increased costs to individual patients of dealing with multiple specialists usually would outweigh the benefit. But we expect to see specialization when the benefit is sufficient. Thus, for example, we see specialization in certain very complex and serious childhood diseases such as cancer.

These simple intuitions go a long way. Divisions in a corporation or agencies in government are like the workers in the pin factory, specializing in particular tasks but limited in their specialization by the problem of coordination. Our problem, however, goes further. We must decide not only how many groupings to have but also which activities to group together. For example, even if we knew that a company should be divided into six divisions, we would still have to decide which activities are to be put into each division. There is, to our knowledge, almost no formal literature on this topic. Instead, the relevant literature dates back to the informal discussions of organizations from the 1970s and earlier. The classic works are the historical studies of businesses by Chandler and the institutional economics of Williamson. The key idea is complementarity, which posits that activities that benefit most from coordination should be grouped together.

Chandler observed that corporations were often originally organized functionally. Functional organization divides the firm into departments in charge of specific functions: sales, production, purchasing, etc. This functional organization was later termed “U-form,” with the U standing for unitary.

The U-form has many advantages. It helps promote coordination and specialization within the functional areas. But Chandler observed that as corporations grew, they discovered several problems with the U-form. One problem was that central management became overloaded with

decisions about daily operations and coordination of the functional units, and it could not focus on strategic decisions. In addition, functional units created agency problems. Employees tended to concentrate on their functional specialties at the expense of the overall profitability of the firm. There was no easy way to monitor employee performance; management could not accurately determine profitability of functional subunits because the subunits did not produce a marketable product on which to base transfer prices.

To resolve these problems, growing corporations tended to reorganize into a divisional structure, with each division organized around a product or product line. This form was later termed the “M-form,” with the M standing for multidivisional. The divisions themselves could be organized along functional or U-form grounds, so that the firm resembled a collection of smaller U-form companies.

The M-form was thought to have solved many of the problems with the U-form. It separated strategic and operational functions so that management could focus on broad strategy, and the operational functions could be coordinated within each division. It also allowed better monitoring of the performance of managers because divisions could be measured by profits, unlike functional pieces.

The government is largely organized along the M-form—that is, by purpose. Congressional committees, executive branch agencies, and departments are like operating divisions. Each provides a “product” to the public, such as tax collection, national security, or education. Like most corporations, the government is not purely M-form. There are a few functional elements, such as a common payroll and pension system, but as a whole, the government resembles an M-form company.

The conclusions from the M-form and U-form literature get us part of the way to an answer. To the extent the M-form is better, we should divide an organization along product lines rather than by function. But the theory does not tell us how to determine the extent of a “product line.” Instead, it assumes that there will be some natural or obvious breakdown of the business into product lines. This, however, is not always the case. Indeed, the scope of a “product line” is precisely the core problem of this Article. For example, it is not clear whether luxury and value hotels are separate products or whether particular types of services offered by pediatricians are separate products. And it is not clear whether various governmental programs should be part of the same “product line” as taxes.

The idea implicit in the literature is that there should be a benefit to grouping activities together. The activities must be complementary. There must be some economy (e.g., economies of scope) to combining them. For example, Chandler quotes an internal memorandum from DuPont, which was then in the process of reorganization:

[please inset the following paragraph]

“The most efficient results are obtained at least expense when we coordinate related effort and segregate unrelated effort. For example, purchase of materials is unrelated to the sale of a finished product in a much greater degree than manufacture and sales, or manufacture and purchasing; and legal work is still more unrelated to either of those before mentioned.”

This intuition is nothing more than the realization that the benefits of coordination of two functions, such as complementarity in performance, can outweigh the benefits of keeping them separate, such as economies of scale or expertise.

Williamson, one of the most prominent proponents of the M-theory, has only a brief discussion of how a company should set up its divisions. In the end, he offers nothing more than the intuition from the DuPont memorandum. In an example, Williamson considers a company that produces five distinct final products. There are three activity stages to producing these products—an early production stage, an intermediate stage in which production is completed, and a marketing stage. He assumes that all products originate in a common first stage. From there, four distinct intermediate stage processes lead to the five distinct final products. This is represented in Figure 1. Note that A<1> and A<2> are the same product, and V<1> and V<2> are the same intermediate production process.

Figure 1. Williamson Divisions

Williamson then discusses how the company should be broken down to divisions. First, he separates the initial stage of production into a division on the theory that at this stage the economies of scale outweigh the benefits of separate production within each division. Then, he combines the intermediate stages and the final stages into a set of divisions on the theory that there needs to be a high degree of coordination between these stages. Product A is split into two divisions on territorial grounds, each combined with an intermediate stage of production. Next, he assumes that there are economies of scope for the intermediate stage of products B and C and, therefore, combines them into a single division with a single intermediate stage. Finally, products D and E are assumed to be complements, so they should be marketed together even though they are produced at the intermediate stage by separate plants. Therefore, he combines D and E along with each of their intermediate stages into a single division.

Williamson offers no clear theory for making these decisions. But his intuitions are similar to those expressed in the DuPont memorandum. Grouping activities together allows coordination of the activities but reduces the benefits of specialization. We want to provide a partition or set of groups that best takes advantage of coordination while minimizing the loss in specialization.

These intuitions can be confirmed in a variety of common contexts. Consider how doctors design specialties. Doctors may specialize in eyes, in feet, or in cancers, but they rarely combine these specialties. It is common, however, for doctors to combine the knowledge of ears, nose, and throat problems into a single practice. How can we explain this? Specializing only in eyes allows doctors to gain the benefits of specialization and coordination of knowledge about eyes with few of the problems of coordinating with other medical knowledge. Specializing in both eyes and feet seems ridiculous because there are few benefits to coordinating these activities or having specialized knowledge in both. But it makes sense to combine ears, nose, and throat practices because they are important to coordinate—problems in one area may be related to problems in other areas. These breakdowns in practice areas, which seem perfectly natural to us, reflect the basic logic of coordination and specialization.

The same is true within businesses. Recall from the hotel example that separation of luxury hotels into a separate division allows specialization in providing luxury hotels and coordination of those types of activities. At the same time, it makes coordination of luxury hotels and value hotels more difficult. The tradeoff is whether the benefits from specialization and coordination within the luxury hotel market are greater than the costs of coordination with other parts of the company's business.

To summarize, the basic tradeoff is an old one between specialization and coordination. Creating more divisions promotes specialization and coordination within divisions but increases coordination costs between them. In deciding how to place tasks within a division or how many divisions to divide a task among, we have to look at the benefits of the various groupings given these costs; the result is that we should group complementary functions together.

B. Departmentalization in a Governmental Context

The departmentalization literature largely focuses on the organization of entities competing in the market. Our problem is slightly different: We want to determine the best organization of a government bureaucracy that is producing and administering policy. This Section develops intuitions about the organization of bureaucracies. As in the market context, the intuitions are based on the benefits of specialization and coordination when different groupings of activities are applied.

The key difference between government and market contexts is that there is no obvious measure to determine how well the government is doing (unlike, say, profits or stock price in the market context). Ultimately, we should measure government output in terms of how well the government produces and implements policy. There is, however, no uniformly agreed-upon metric for measuring the effectiveness of government policy.

Because our question is about bureaucratic organization, we can assume that the basic policy that we are considering is fixed, having been set by the principal (the executive or the legislature). We can then measure how well the bureaucracy implements that policy by measuring how accurate the implementation is for a given cost. A more accurate policy better distinguishes between different individuals or different actions. It comes closer to the optimum. Accuracy, however, is expensive, which means that no policy will be implemented with perfect accuracy.

There are a number of prior works that discuss the tradeoff between accuracy and the costs of producing accuracy, commonly labeled complexity. For our purposes, we can simply think of the problem as trading off marginal benefits and marginal costs. For example, a single speed limit for all roads would be highly inaccurate. We can increase accuracy by posting separate speed limits for different roads, although this increases costs. We could make the speed limit policy more accurate by posting different speed limits for individual roads under various weather and traffic conditions, but this would further increase costs. At some point, the marginal cost is not worth the marginal benefit of the increase in accuracy.

This approach, however, considers only the costs and benefits of different levels of accuracy within a given institution. The issue presented here is slightly different—namely, the assignment of policies across different agencies. Because the costs of producing accurate policy will vary with

institutional structure, the independent variable is not how much accuracy to produce given some cost of production. Instead, the independent variable is the institutional structure itself, which then determines the tradeoff between accuracy and complexity.

Given this setup, the analysis works basically the same way that it does in the private sector. Grouping activities together allows coordination of those activities but reduces coordination of those activities with others. Smaller groupings allow more specialization; larger groupings, more coordination. The key is to group activities that are related—that is, there are large benefits to coordination and low costs to the loss of specialization.

For example, consider the IRS and the Department of Agriculture. The IRS can presumably measure income at a given level of accuracy better than the Department of Agriculture, because of specialization by the IRS in income measurement. Similarly, the Department of Agriculture can presumably best measure items related to agriculture, such as the value of various farming techniques or the safety of various methods of preparing food. Separating these activities allows this specialization, but also creates coordination problems. While the lack of coordination between these activities will look like bad policymaking, it may in fact be optimal: Setting up the bureaucracy in a way that instead coordinated the activities of the USDA and the IRS would reduce the benefit of each of these agencies specializing in its own activities.

Analyses that look only at one aspect of the problem rather than overall government policy can be faulty. For example, Victor Thuronyi notes:

The dairy farmer subsidies include accelerated depreciation deductions on livestock and equipment and the acceptance of “cash accounting,” both of which defer tax liability with no interest. While these tax provisions subsidize production and encourage herd expansion, the Department of Agriculture, on the other hand, pays dairy farmers to curtail production and slaughter their herds.

Thuronyi presents this clash of policies as if it is necessarily a bad idea. It may, however, be the result of the best possible choice of organizational form. The IRS might be the best agency for providing investment subsidies, and the Department of Agriculture is probably the best agency for regulating farmers. Certainly, neither agency would seem to be best at doing both activities. Therefore separate agencies for each function, and the resulting lack of coordination, could be optimal. Any other organizational form, including one that coordinated these conflicting policies, might be worse. It is not that we should applaud the end result as directly desirable; indeed, this particular example may be the result of a failing of process or design rather than a result of optimal organizational structure. We need to recognize, however, that desirable separation of functions into divisions is going to lead to lack of coordination. An organizational form that produces better coordination may very well be inferior on other grounds and, therefore, the lack of coordination may be optimal.

This analysis flips the usual complaint about tax expenditures on its head. Many claim that tax expenditures increase the complexity of the tax system. Under the analysis here, however, a decision to have separate tax and spending programs is a decision to have specialization in each program, presumably with more accurate and detailed measurements within each program—i.e.,

greater complexity. A decision to consolidate spending programs into the tax system, by contrast, is a decision to coordinate the programs with less fine-tuned accuracy in each program. Therefore, putting a program into the tax system can be seen as a decision for simplicity.

To see this, consider another example. Suppose we are considering whether parts of the welfare system should be integrated into the tax system, as the EITC is. We must compare the benefits of having two programs and two administrative agencies (a special agency to administer the welfare part of fiscal policy and the tax agency to administer the tax part) to the benefits of having a single agency administering both programs.

If we have separate programs, each program can be more easily tailored to meet its specific goals. For example, if it is desirable to have monthly accounting periods for welfare and annual accounting periods for tax, each program can adopt the desired period. Similarly, if the welfare system requires one measurement of “need” and the tax system optimally uses a different measurement of “ability to pay,” each program can adopt the required definition. Separation may enable administrative specialization in the specific requirements of each policy, and hence improve its accuracy.

The disadvantage of separate agencies is that the various welfare programs and the tax system may not be coordinated very well. For example, welfare programs typically contain phase-outs, which act as a marginal tax on income. Failing to coordinate these phase-outs can lead to very high effective marginal rates and a marginal rate structure that seems random.

In addition, each welfare program might use its own eligibility test, which would mean that individuals would end up providing similar but slightly different information to various government agencies. For example, the same child might qualify as a family member under one program but not under another, or various elements of income might be included in one program but not in another, creating enormous complexity for individuals.

We can generalize this example. Which integrated transfer programs are likely to be successful? Programs where the coordination benefits between the tax system and the other program are high and the specialization benefits of separate programs are low. Thus, we want to integrate programs that have close complementarities with the tax system—e.g., programs that measure along similar margins. This is why welfare and tax are such an obvious pairing. They measure along very similar boundaries.

One scholar, Eric Toder, recently provided a list of the features that make a program a good or bad fit for administration by the IRS. Several items on this list relate to budgetary aspects of the problem, which we do not deal with here. Most of the remaining items relate to the benefits of coordination and specialization and can be explained by our framework. For example, Toder argues that if the agency has a high degree of discretion in setting policies, implementing the program through the tax system may be unwise. The reason must be that if the agency has discretion, it will be using expertise to make determinations, which means that the value of specialization is high. Toder also argues that the more that the spending program uses tax return data for eligibility, the more desirable integration is. This easily fits within our framework—the IRS has expertise in measuring along those margins, and it exhibits economies of scope in such measurement. Toder

argues that programs that have open-ended eligibility are better suited to the tax system than other programs. The reason is that the tax system uses end-of-year filing, so that programs requiring up-front eligibility testing would need specialization that would not be complementary with that of the IRS. Toder's suggestions seem eminently sensible within our framework.

The problem with these intuitions is that the terms are extremely vague and are at a very high level of generality. Translating these terms into measurable formula for making decisions is far from an easy task. But relatively crude ideas about accuracy, complexity, specialization, and coordination can help policymakers muddle through the problems they face. To see this, in the next Part we go through two examples in detail.

IV. Applications: The Earned Income Tax Credit and the Food Stamp Program

This Part applies the framework developed above to two major welfare programs—the Earned Income Tax Credit and the food stamp program (FSP). The welfare system is of special interest in the analysis of the integration question. In 1998, about \$400 billion was spent on more than eighty means-tested programs in the United States. Total spending on cash and in-kind welfare benefits was more than five times higher in 1998 than in 1968 (adjusted for inflation), while the U.S. population increased 35% during the same period. The share of the federal budget used for means-tested programs rose from 6.4% in 1968 to 16.8% in 1998. In addition, the composition of tax expenditures changed significantly over the last two decades of the twentieth century. Social tax expenditures as a percentage of GDP increased over 40% during this period while business tax expenditures were cut in half. Social tax expenditures accounted for 79% of all tax expenditures in 1999, compared to 57% in 1980. These trends underscore the importance and relevance of the integration problem to welfare reform.

We choose to compare the EITC to the FSP for three reasons. First, the programs are to some extent similar, yet one is integrated into the tax system and one is not. Comparison of the performance of the two programs gives us some sense of the costs and benefits of integration.

Second, plausible cases can be made for integrating both programs with the tax system primarily because their eligibility criteria are income-based. In addition, there are serious problems with integrating each of the programs with the tax system. Therefore, they make for interesting programs to study.

Finally, these two programs are among the largest welfare programs in the United States. Accordingly (and perhaps also for the other two reasons discussed above), there have been a large number of studies on the delivery of these sorts of benefits through the tax system, so rather than building from scratch, we can analyze the conclusions of these studies within our framework. Much of the earlier work on combining tax and transfer systems concerned the negative income tax, with the basic claim being that benefits can be distributed more efficiently through the tax system. We agree with the idea that some benefits can best be distributed through the tax system and, in particular, believe that provision of the EITC through the tax system makes sense. But we argue, contrary to the thrust of the negative income tax literature, that some welfare policies are best implemented separately because of institutional considerations.

Our analysis follows in part Anne Alstott's work on the EITC. She is critical of arguments in favor of integrating tax and transfer systems, arguing that the tax system cannot adequately perform many functions of the transfer system. For example, she argues that the tax system cannot respond as quickly to emergencies as the transfer system. Alstott's arguments can be analyzed within our framework and are largely consistent with it. Our conclusions in many places are different from hers, but to a great extent the particular conclusions are less important than the mode of analysis. Therefore, we use her work as a focus of our discussion.

We begin with background information on each of the programs and discuss the reasons for integrating the tax and transfer systems. We then turn to the details of each program and conclude with a brief discussion of negative income taxes more generally.

A. Background

The FSP is an in-kind transfer program jointly administered by the USDA and state agencies. It provides food assistance to families and individuals based on monthly income, assets, and family structure. Monthly income is effectively equivalent to cash income, subject to several exemptions and deductions. To qualify for the FSP, the applicant's monthly income must be below a threshold that varies with family size and composition. Household assets typically cannot exceed \$2000, excluding certain assets (e.g., a house and lot, vehicles of specific use or of limited value, and most retirement pension plans). Benefit levels vary according to family size and composition, such as whether a family member is elderly or disabled; they are calculated by taking the maximum benefit for a household of a given size and composition and subtracting thirty percent of its net income, effectively imposing a thirty percent tax rate on beneficiaries.

Eligibility for food stamps is determined in advance. To apply for benefits, food stamp applicants must visit a state office in person during regular business hours. Applicants must fill out a detailed and often lengthy application and provide supporting documentation. Over forty percent of food stamp applicants make two or more trips to the state office to complete the initial application process. Eligibility must be recertified throughout the year, often on a monthly basis and often requiring return visits to the state offices.

The EITC is a wage subsidy implemented as a refundable tax credit. As wages rise to specified levels, the credit increases, eventually plateaus, and then phases out. The levels of the phase-ins and phase-outs depend on family size, with larger families generally receiving more generous benefits. Participants claim eligibility on their tax return without any prior certification process. Participants are, however, subject to ex post audits through both data matching and other auditing techniques. Typically, benefits are paid once a year after the individual files a tax return. To the extent the credit offsets taxes otherwise due, individuals can adjust their wage withholding to get some of the benefits during the year. In addition, the credit can be obtained during the year through an advance payment system, although, as discussed below, this option is rarely used.

The argument for the integration of transfer programs such as the EITC and the FSP into the tax system is that integration enhances administrative efficiency by reducing bureaucratic costs and complexity. In the language of the framework proposed here, there are benefits to specialization and to coordination from integration.

The benefits to coordination arise because the FSP and the EITC (and, for that matter, virtually all means-tested welfare programs) rely on similar criteria, such as income. The IRS is specialized in measuring these criteria and is therefore likely to be able to take these measurements more accurately and more cheaply than other agencies, such as the USDA. In addition, as a form of coordination, integration would save the costs of processing the same or similar information more than once by making use of a single preexisting measurement system. Integration would also confer savings on recipients, eliminating the need to fill out multiple forms and visit multiple government offices.

Further benefits to coordination arise through the interaction of tax and transfer programs on a variety of margins. For example, phase-outs of means-tested programs increase effective marginal tax rates, and failure to coordinate phase-outs can create extremely high rates for low-income individuals. Similarly, the amount of benefits for any program should depend on amounts provided by other programs. A family receiving the EITC might, as a result, have less need for food stamps or vice versa. Definitions of marriage and family also need to be coordinated. Both the tax law and welfare programs can create marriage penalties or bonuses. Coordination of these programs to limit the effects on marriage may therefore be important.

Empirical evidence on the accuracy and complexity of the EITC and the FSP further support the argument for integration. Subject to data and conceptual limitations, the key facts are that the FSP costs more to implement than the EITC by an order of magnitude, but the FSP is not any more accurate than the EITC (although it produces a very different pattern of errors).

In particular, annual administrative costs of the FSP are estimated to be about \$4 billion per year. Individual compliance costs are also likely to be high because of the certification, recertification, and reporting requirements. There are no firm estimates for the administrative costs of implementing the EITC because the IRS does not separately account for the cost of administering the program. About ninety-five percent of EITC claimants would file tax returns even if not eligible for the EITC, which means that the IRS would process most of the information and individuals would incur the cost of filing returns anyway. Individuals must file an extra form with their return to claim the EITC, which increases their costs somewhat. The EITC also creates additional audits, which impose costs on both the government and individuals. In light of these facts, estimates have put the administrative costs of the EITC at about \$ 320 million, or less than one-tenth the amount spent on the FSP. Even at this lower cost, the EITC is substantially larger than the FSP.

Notwithstanding the vastly higher administrative and compliance costs of the FSP, it is not clear that it is any more accurate. The FSP has a lower error rate than the EITC, but also a much lower participation rate. EITC overpayments are around 27%, while FSP overpayments are around 7%. Underpayment rates are similar between the two programs: EITC underpayments are below 2.5%, and the FSP's are around 2.9%. On the other hand, the EITC participation rate is about 89%, while FSP participation is around 70%. Combining each program's underparticipation rate with its underpayment rate to determine an overall underprovision rate, we get the following picture:

Table 1. Comparison of the EITC and FSP

We may not weigh these types of errors equally, so we cannot say for certain how these numbers compare, but we see no overwhelming reason why the FSP error rate should be preferable (and we argue below that in fact the EITC error rate is actually preferable). Indeed, both the USDA and the GAO, in analyzing the FSP, weight the two types of errors equally. Therefore, when accuracy is measured based on under- and overprovision, the FSP, while generating a different type of error than the EITC, cannot be said to be more accurate than the EITC, even though it costs ten times as much to administer and is only one-half the size.

The question posed is whether a change in institutional structure can improve the performance of either the EITC or the FSP. The argument here is similar to that invoked to support a negative income tax: The idea is that a single approach to need or to well-being should be used in the tax and transfer system. A consistent methodology makes sense from a policy perspective—need is need, and there is no reason to change the approach when a tax goes from positive to negative. There is nothing special about zero. In addition, there are great institutional advantages to a single system, such as the savings of administrative and compliance costs. The concern, addressed in the next Section, is that integration would undermine the benefits of specialization that come from addressing the particular administrative problems presented by food stamps.

B. Food Stamps

Given these reasons for integrating the FSP into the tax system, one might think that the answer is clear. There are, however, a number of aspects of the FSP that are not complementary to the tax system. Integration might involve compromises to these aspects of program design.

There are four areas of concern, three of them found in Alstott's discussion of the EITC, and a fourth not applicable to the EITC, and therefore not discussed by Alstott. First, there are important differences in eligibility measurements and criteria between the tax system and the FSP. Arguably, an ideal transfer system might use different measurements from an ideal tax system. If the systems were integrated, they would use a single set of criteria, so that integrating the two would involve a loss of accuracy. Second, the pattern of errors between compliance and participation is different in the two systems. Integrating the FSP into the tax system would likely substitute the tax pattern of errors for the FSP pattern of errors. Third, the tax system tends to use long measurement periods to measure permanent changes in well-being, while welfare systems use short prospective measurement periods to be responsive to temporary changes in circumstances. The fourth concern, not mentioned by Alstott because she focuses on the EITC, is that food stamps are provided in-kind while the tax system is specialized in collecting and distributing cash.

We review these considerations below, showing how they relate to issues of specialization and coordination, and how they fit within the general framework discussed above. Each of these potential areas of concern can be seen as a problem of accuracy. The question is whether a specialized food stamp program is necessary to provide that accuracy or whether coordination with the tax system and the use of the tax system's resources—which are specialized to handle income measurement—is superior. Our conclusions differ somewhat from Alstott's conclusion regarding similar issues with respect to the EITC. In particular, we conclude that only the responsiveness problem—the problem of measurement periods—presents an obstacle to integration of food stamps into the tax system. The data on responsiveness are mixed, and we are uncertain of what the costs

to responsiveness from integration would be, but there is at least a significant chance that the responsiveness problem is sufficiently serious to merit caution in integrating the programs. We do not want to put too much stress on these differences, however. Instead, our goal is to show how the analysis fits directly into the institutional framework we posed above.

1. Measurement Criteria

On the surface, the tax system and the FSP (as well as almost all means-tested transfer programs) use similar eligibility criteria—usually income, assets, and family composition (including whether a family member has a special status, such as disability or advanced age). This commonality is the primary reason that integration of welfare policy seems potentially attractive. Nevertheless, as Alstott notes, there are important differences.

In particular, Alstott argues that the income tax system tends to use a less accurate measurement of income and other resources than traditional transfer programs. Most transfer programs attempt to measure all sources of cash available to meet the living expenses of a family, with limited deductions. For example, the eligibility criteria for food stamps look to all cash income, allowing deductions (other than standard deductions and a fixed deduction against earned income) mainly for dependent care, medical expenses, and excess shelter costs. The tax system, however, excludes many sources of wealth, such as retirement benefits and interest on state and local bonds.

Transfer programs also tend to measure the value of recipients' assets, which is not generally done in the tax system. Thus, food stamps limit eligibility to families with less than \$2000 in certain assets, with some adjustments for vehicles. By contrast, the EITC does not have an asset or wealth test.

The definition of a family is likewise more nuanced under welfare programs than under the tax system. Current tax law has a very narrow definition of family, generally treating even children in the same household as separate taxpayers. Income transfer programs, in contrast, adopt a broader definition of family that includes most individuals who live together. For example, the FSP defines a household as a group of people who live together and customarily purchase food and prepare meals together.

Adopting these tax definitions might create problems for a transfer system. As Alstott notes, in the 1960s and 1970s, when the negative income tax was receiving serious academic attention, scholars argued that the income definition in the tax law was inadequate for transfer programs. The tax system's measurement of income is sufficiently crude that some families have received the EITC even though their incomes, if measured using a comprehensive base, could be as much as \$ 75,000.

This conclusion argues against integration. One of the main advantages to integrating transfer programs such as the FSP into the tax system is to use common definitions and infrastructure. In our framework, integration can take advantage of the specialized income measurement resources of the tax system and gain the benefits of coordination with the tax system. But using tax definitions might result in a significant loss of accuracy for the FSP.

There are several responses to these arguments. First, it is not clear why transfer systems and the tax system use different definitions. Alstott merely observes that they do and argues that integration would involve a compromise of current practices. But to determine the social cost of such a change, we need to understand why the different programs use different definitions in the first place. There is nothing special about welfare programs that should make us want to use different definitions: It is not clear why we should care more about accuracy when giving very small sums as transfers to a poor individual than when taking thousands or millions in taxes from a very wealthy individual.

Second, assuming that the tax system is inaccurate, we must weigh the value of accuracy against its cost. If the savings from integration are substantial, it may be worth the loss in accuracy. Merely noting that there might be a loss in accuracy is insufficient.

Third, even if transfer programs use more accurate definitions of income, they are unlikely to be very good at measurement. In fact, notwithstanding attempts to measure well-being carefully in the FSP, about half of all FSP payment errors are due to an incorrect determination of the household's income. The tax system, on the other hand, specializes in income measurement and has invested billions of dollars in this expertise, including in computer systems that can match income reports and in agents trained to ferret out fraud. There are, of course, substantial mismeasurements by the tax system, but our guess is that for any given measure of income, the specialization of the tax system is likely to result in better measurement per dollar spent. It is not clear that the attempt at higher accuracy in transfer systems such as the FSP actually succeeds.

Finally, even if different definitions of well-being are necessary for food stamps and taxes, the IRS might still be best at measuring along those margins. That is, merely because we are going to have the complexity of different definitions of well-being for different programs does not mean that integration does not make sense. For example, the FSP measures wealth to determine eligibility while the EITC does not. Under FSP rules, however, the wealth measurement is largely limited to cash-like assets. These assets—with a clear market value—can be easily measured by the tax system. The IRS might therefore be the best agency to measure each criterion; all that would be required would be to alter the definitions of the tax system to achieve an acceptable level of accuracy. The problem of different measurement criteria is thus not a good argument against integrating the FSP into the tax system.

2. Error Patterns

As noted, food stamps have a low participation rate but also a low overpayment rate. Both facts are likely attributable to the eligibility process. Precertification is very costly for low-income individuals, which discourages FSP applicants but at the same time weeds out false claims. The EITC has a high participation rate but also a high overpayment rate. These facts are likely due to the lack of a precertification process. Individuals need merely to file a tax return to claim the EITC. Most individuals claiming the EITC must file a return anyway, so the additional cost of claiming the EITC is relatively low.

Alstott refers to this pattern of errors as a tradeoff between participation and compliance. The FSP chooses to have a high compliance rate at the cost of a low participation rate while the EITC opts

for a high participation rate at the cost of a low compliance rate. Given that information is costly, such a tradeoff between participation and compliance seems inevitable.

Alstott takes the existing pattern of welfare programs such as the FSP (low participation, low overpayment) to be desirable. We are not sure we would agree. It is not at all obvious, in light of the tradeoff, which pattern is preferable. As noted above, both the GAO and the USDA weight under- and overpayments equally.

Although there is no general theory on how to balance under- and overpayments, we can identify three relevant factors. First, how we choose the balance may depend on which individuals fail to receive payments for which they are eligible and which individuals get overpayments. For example, Liebman examines which individuals get EITC overpayments and finds that many overpayments go to families that are quite similar to eligible families. Ineligibility often results from a failure to meet one or more minor requirements rather than because overall income is too high. In such cases, we may not be concerned by overpayments.

Second, overpayments require higher tax rates, leading to excess burden. We would need to know the marginal cost of funds to determine the size of this effect. Finally, for any given belief about the tradeoff between overpayments and underpayments, we must examine the administrative cost of reaching the desired balance between the two. It will generally be administratively cheaper to have more overpayments but full participation than to prevent overpayments at the cost of less participation. The reason is simple: It takes resources to prevent overpayments.

While the tradeoff between these factors may be complex, if we examine the actual numbers the FSP pattern does not look attractive compared to the EITC pattern, given almost any imaginable views on how to balance these factors. In particular, Janet Holtzblatt reports that if we add overpayments and administrative costs together, the EITC and the FSP both incur costs of about 25 cents on the dollar. This means that, as compared to the EITC (and by extension, the FSP if it were administered through the tax system), it costs the current FSP a dollar of administrative costs to prevent a dollar of overpayments. This cannot be justified. For example, suppose that a welfare program has 100 truly eligible recipients who each receive \$1, and that the tax system creates no administrative costs but has overpayments of 25 cents on the dollar while the FSP has 25 cents on the dollar of administrative costs and no overpayments. The cost of both the FSP and the tax system would be \$125. It is hard to imagine that we would want to essentially throw away the \$25 on administrative costs rather than transfer it to individuals.

Moreover, these numbers may significantly understate the problem. Holtzblatt's study, from which these figures are derived, does not consider compliance costs of participants: As a society, we thus spend more than a dollar to prevent a dollar of overpayments. Holtzblatt also does not consider the social costs of FSP underprovision due to nonparticipation. Once these two additional costs are factored in, the price of preventing overpayments for the FSP seems exorbitant. Finally, factor in the data showing that overpayments under the EITC would go to families that are similar to eligible families. If overpayments under an integrated FSP were similar, there would seem to be no argument for the FSP pattern of errors. Therefore we can conclude that the pattern of errors for the EITC is preferable. The error pattern argument supports integration.

3. In-Kind Provision of Benefits

The FSP provides benefits in-kind through a voucher-type system. The government does not directly provide food to individuals on food stamps, unlike, say, those receiving subsidized school lunches. Instead, the government provides individuals with coupons (typically in the form of an electronic card) that can be used only to purchase food at USDA-certified establishments. Authorized retailers must fill out an application showing that they meet the requirements and personally go through a training session. Integrating the FSP into the tax system would probably involve eliminating this in-kind delivery, because the tax system could not easily engage in this certification.

The rationale for in-kind transfers has been subject to extensive study. While the merits of such transfers have been considered in detail, arguments made in these studies might not apply to transfers of food (as opposed to insurance or durable goods). Fortunately, we need not delve into that literature here. The reason why is that the typical food stamp benefit is less than the monthly food budget for most beneficiaries, and is only slightly restricted in its approved food items. In theory, therefore, the program's decision to provide in-kind benefits or cash assistance should have no effect for these beneficiaries; the entire food stamp amount will be spent on food either way. This prediction is generally borne out by observed evidence, although there are small and persistent deviations. The reasons for such deviations are not fully understood.

Although some may value the increase (if any) in food consumption due to in-kind provision, it is doubtful that any such benefit is worth the cost. If we measure nutritional outcomes as opposed to food intake, we cannot say, based on available evidence, that providing cash assistance rather than food stamps would have any effect. And it is nutrition, not caloric intake, that we should be concerned about with a food program: As the Food Stamp Act itself declares, food stamps are aimed at "raising levels of nutrition among low-income households."

"Cashing out" food stamps would also save administrative and compliance costs above and beyond the immediate savings from integrating food stamps into the tax system. An individual wanting to use the value of food stamps for something other than food has an incentive to sell them for cash. This means that the FSP must spend resources to monitor and enforce the use of food stamps, and that individuals (and food retailers) will incur costs to comply with the food stamp police. Some of the costs of enforcement have gone down in recent years with the switch to electronic benefits, which are harder to trade, but the problem is still significant.

Finally, European welfare systems do not have specific food transfer programs. There is no evidence that their absence leads to systematic effects on food security or nutrition.

One argument against cashing out food stamps through integration with the tax system is that it would be too big a change in the purposes of the program. Cashing out the FSP would be like eliminating it and increasing the EITC or TANF by a comparable amount. The argument of this Article is that we should compare institutional performance in meeting previously defined goals. The goal of food stamps might be defined as improving nutrition among the very poor, in which case cashing out would seem to have little effect on the achievement of such a goal. But the FSP's goal might also be to provide food to the poor. It is not clear why we would opt for a goal of

providing food over providing nutrition, but the integration decision must take as given the basic policy goals of the program, and it is not inconceivable that providing food itself is the goal. If so, integration should not change those goals.

Providing food itself (rather than money to be used for food) is obviously not complementary to the services provided by the tax system, which is why we assumed that integrating food stamps into the tax system would mean cashing them out. If the FSP provided food itself rather than vouchers for food, integration with the tax system would probably be completely infeasible. For example, it is hard to imagine the tax system directly providing school lunches. But the FSP uses a voucher system combined with a system for certifying food providers. This structure might be susceptible to a hybrid system in which the tax system provides credits or other benefits for expenditures on food at authorized establishments. A separate agency that has expertise in certifying and monitoring food providers (such as the USDA) could perform the certification function. The Hope and Lifetime Learning tax credits, for example, work exactly this way.

Such a hybrid system may not be easy to implement. Individuals would have to keep receipts to prove their expenditures, fraud would be common, and coordinating between the IRS and the other agency might be difficult. Nevertheless, given the enormous costs of implementing food stamps as is, such a change may be beneficial even if it is expensive.

In sum, if the FSP were to be integrated into the tax system, cashing out food stamps would be the easiest course. The tax system could much more easily provide cash benefits than in-kind benefits, because it specializes in measuring income and collecting or disbursing cash. Cashing out food stamps would probably not reduce the nutritional benefits of the program and would reduce or eliminate enforcement of its in-kind limits. If the purposes of the FSP cannot be served by cashing out the program, integration may still make sense through a hybrid-type system, but it would have many fewer benefits.

4. Measurement Periods and Responsiveness

Responsiveness is by far the most difficult problem with integrating the FSP with the tax system, and we think that it explains why food stamps are not, and perhaps should not be, integrated into the tax system. The problem is that food stamps are provided on a monthly basis and eligibility is determined on a similar time scale, while the tax system operates on a yearly accounting period. Alstott describes this difference as a tradeoff between responding immediately to short-term changes in need and waiting to evaluate the permanency of the changes. Welfare systems such as the FSP tend to respond to short-term changes, acting in some ways as insurance, while the tax system tends to try to measure longer-term changes in well-being.

Responsiveness can be very important to a welfare system. If an individual or family has a short-term but severe drop in income, they can lose housing, go hungry, or otherwise go without basic needs. The effects can be large even if the drop in income is only short-term. In theory, if the drop is only short-term, capital markets (and insurance) could help smooth consumption, but in practice, capital and insurance markets are highly imperfect. Moreover, if the immediate crisis turns out to be a permanent change, capital markets may not be able to help even if responding immediately to the permanent change is vital. Responsiveness is an important function for welfare systems.

The tax system does not value fast responses to changes in well-being. The tax system redistributes income from the wealthy to the poor, thereby increasing the welfare of the poor. In this sense, it is complementary to welfare systems. But the tax system does not generally attempt to provide emergency assistance. Instead, as Alstott notes, the tax system tries to measure permanency of changes in well-being. In fact, the one-year accounting period is arguably too short to measure long-term well-being, and scholars have proposed averaging systems that remedy this shortfall.

Integrating food stamps into the tax system, therefore, might mean that it would not be responsive to short-term needs, seriously undermining its performance. In the language of specialization, the FSP specializes in measuring short periods while the tax system specializes in long periods. Integrating the FSP into the tax system would mean losing the benefit of specialized FSP measurements. To understand this problem, we explore three issues. First, we try to assess how important a role responsiveness plays in the FSP. We would like to know whether reducing the responsiveness of the FSP would be a significant problem by itself. Second, we would like to know whether there are good substitutes for the short-term needs fulfilled by food stamps, so that there need be few social losses if the FSP no longer serves this role. Third, we would like to understand how responsive the FSP could be if it were integrated into the tax system.

a. Emergency Responsiveness Appears To Be Moderately Important to the FSP, but Not Central

Responsiveness is built into the design of the FSP. The law provides that households are entitled to apply for food stamps on the first day they contact a food stamp office during business hours. States must provide eligible applicants with food stamps as soon as possible, but no later than thirty calendar days following the date the application was filed. In addition, households that have very little income or liquid resources can apply for expedited food stamps. If eligible, an applicant must receive them no later than seven days after the date of application.

There have been lawsuits about whether states are meeting these requirements. These suits may merely represent transfers—individuals might sue to speed up the process because sooner is better. But they may also represent real need created by delays in processing. For example, scholars have found that about two-thirds of all people entering the FSP have experienced a twenty percent drop in income in the prior four months. In reality, the suits probably represent some balance between these two motivations, and indicate that a significant concern in program design is to get food stamps to needy individuals quickly.

Another indicator of the importance of responsiveness is the duration of spells in the FSP. If individuals or families generally have short spells, it would indicate that the system must be responsive—the system would be helping with short-term needs and failing to be responsive would defeat that purpose. If, on the other hand, spells are long, the tax system might be able to help because the program would be responding to more permanent changes in well-being.

The data on spells are mixed. Philip Gleason, Peter Schochet, and Robert Moffitt have done the most extensive study of food stamp spells. They examine two samples—one comprised of people who begin a spell of FSP participation during a given calendar period (“entrants”), and the other comprised of all food stamp participants at a given point in time, regardless of when they began

receiving food stamps (“cross-sectional”). They report that most entrants exit fairly quickly, with a median participation period of about nine months. Furthermore, fewer than one-third of entrants remain on food stamps for two or more years. Nevertheless, the cross-sectional sample shows that at any given time, individuals on food stamps are likely to have long spells. These two samples may lead to the conclusion that a certain group of FSP recipients goes on and off the program repeatedly. Indeed, they estimate that two-thirds of all FSP entrants are repeat entrants, and that many reenter rather quickly after they have exited the program. Accordingly, if short food stamp spells by individuals who go on and off food stamps are aggregated, the vast majority of spells are relatively long-term.

An earlier study by Rebecca Blank and Patricia Ruggles shows similar findings. It reports that much food stamp participation among women occurs during long spells of eligibility. While there are a relatively large number of short eligibility spells for food stamps, most of these spells open and close without program participation. In essence, Blank and Ruggles find that there are two groups of women among those eligible for food stamps: “A relatively disadvantaged group with low future income expectations who enroll in public assistance immediately, and another less disadvantaged group who (largely correctly) predict future income increases and who do not seek benefits.” This suggests that food stamps serve longer-term needs that might be served by the tax system.

One might be tempted to conclude from this data that responsiveness is not important to food stamps. But Blank and Ruggles report that those who take up food stamps do so almost immediately upon eligibility. The tax system, even if able to handle these individuals’ long-term needs, might not be able to respond quickly when the individuals first become needy.

If the primary concern is quick responses to long-term needs, we must understand the immediate cause of the problem and the type of population it affects. For example, it is important to know the source of the 20% reduction in income that often triggers a food stamp application. Only 27% of food stamp participants earn their income through wages; the rest have other sources of income, such as SSI (31.7%), TANF (25.8%), and Social Security (25.5%). If, for example, the reduction in income is mainly due to decreases in earned income, an increase in the size or length of unemployment insurance may be the appropriate short-term response. The little evidence we have found suggests that though only 27% of FSP participants earn income, a decrease in earned income is the most probable trigger for entry.

In addition, retaining the entire food stamp program to solve the immediate needs of a small portion of the population on food stamps may not be desirable. The USDA reports that of 7335 surveyed households receiving food stamps in 2000, 401 were entrants. This means that responsiveness during the year is important only to about 5.5% of beneficiaries. In addition, in fiscal year 2000, only 6.2% of food stamp recipients used the expedited food stamps option, and about 30% of them were new entrants. We may be better off responding to these emergencies in a different way if the benefits from integrating food stamps with the tax system are otherwise large.

In sum, we would like to have more data before claiming that responsiveness is central to food stamps. For example, we would like further study of the length of food stamp participation spells.

Additionally, we would like to understand better the reasons households go on food stamps, and what resources they have that could help them to respond to short-term need.

b. Other Programs, as Currently Constituted, Could Not Adequately Replace Food Stamps

There is a large system in the United States designed to respond to short-term dire food needs known as the Emergency Food Assistance System (EFAS). The EFAS consists largely of food kitchens and food pantries (and their wholesale suppliers). Individuals in need can obtain food from either of these sources. The EFAS, however, is only about one-fifth of the size of the FSP and is not as well-targeted. Moreover, the EFAS is not equipped to handle mid-term problems—the individual who needs more than a few nights of assistance but less than several years—which the FSP handles well. The EFAS could be expanded and modified to some extent were food stamps integrated into the tax system, but this would effectively mean less than full integration because some of the functions of food stamps would be transferred to the emergency food programs.

There might be non-food-related responses to short-term need. If, as suggested earlier, cashing out the program would not present serious difficulties, responsiveness needs might be met through quick cash assistance rather than through quick food assistance. Determining that responsiveness is essential to the FSP would then mean determining whether existing cash programs (combined with the EFAS) could meet these needs.

c. The Tax System Could Not Easily Be Made Responsive to Short-Term Food Needs

We can divide tax system responsiveness into two pieces: the measurement period and the period in which benefits are provided. These can vary. For example, benefits can be provided every month even if eligibility measurements are made only once a year.

The tax system probably could not provide short measurement periods. As noted, it is important for the tax system to use long measurement periods. Reducing the measurement periods in the tax system to allow integration of the FSP and other transfer programs would impose costs on everyone, because shorter periods would require more frequent filings. This could be prohibitively costly.

It is not clear how important short measurement periods are. Short measurement periods, even if otherwise desirable, increase administrative and compliance costs. To the extent short measurement periods increase compliance costs, they reduce participation in the program. Thus, short measurement periods may increase responsiveness for some at the cost of completely eliminating benefits for others.

In fact, because of the difficulties of balancing responsiveness and long-term measurement periods, and because of its low participation rate (due to the high administrative costs of short measurement periods), the FSP has been moving toward longer measurement periods. Currently, the frequency of measurement varies across states and can be quarterly, semiannual, annual, or in some cases biannual. Thus, shifting to annual measurement for food stamps would be consistent with the trend of the program.

Monthly provision of benefits is a more difficult problem. It is not inconceivable that the tax system could provide monthly benefits. The EITC is currently available on this basis under the “advance EITC.” The EITC program, however, has a natural intermediary between the IRS and the needy individual: the employer. The individual certifies to her employer that she expects to be eligible for the EITC and receives installment payments of the credit each pay period. The employer advances funds to the individual and receives a credit on its tax return, effectively acting as a financial intermediary. The FSP, however, has no similar intermediary because individuals need not be employed to receive food stamps. Conceivably banks or other financial institutions could fulfill this role, but this would probably require significant monitoring, potentially defeating the benefits of integration.

Moreover, the advance EITC, even though easier to imagine than an advance food credit type system, has not worked. The take-up rates for the advance EITC are extremely low. Only 1.1% of EITC recipients used the advance payment option in 1998. We suspect that a similar system for food stamps would be even less successful.

d. Conclusion on Responsiveness

Combining these considerations, we conclude that the tax system could not be adequately responsive to the needs that the FSP fulfills. In particular, to the extent that immediate response is necessary even for individuals with medium-to long-term needs, the tax system may be inadequate. Becoming more responsive would sacrifice too many goals of the tax system, and not becoming more responsive would sacrifice too many goals of the food stamp program.

Nevertheless, we think the issue should be analyzed further. The data on the need for quick responses by the FSP are mixed, and we do not yet know exactly how important that need is. Moreover, we do not know whether other elements of the transfer system, if enhanced, could satisfy this need while allowing us to realize the benefits of integration.

5. Conclusion

There would be many benefits to integrating the FSP into the tax system. The two programs have strong complementarities. Integrating the FSP into the tax system would allow the FSP to take advantage of the tax system’s specialization in income measurement, including its filing, computer, and audit systems. Moreover, integrating the FSP into the tax system would allow better coordination of the tax system and the welfare system. A crude measure of accuracy reveals no difference between the FSP and EITC, even though the former is at least ten times more expensive to administer (and possibly much more). Nevertheless, if alleviating short-term hunger is an important enough goal of the FSP, integration is probably not desirable because the tax system is not sufficiently specialized in solving short-term problems. The loss in accuracy from integration would outweigh any corresponding benefits.

C. The Earned Income Tax Credit

The analysis of the EITC is very similar to the analysis of food stamps. The main differences are that (1) the EITC is a cash rather than an in-kind program, so concerns about providing cash

benefits are eliminated, and (2) responsiveness is less important for the EITC than for food stamps. Therefore, we conclude, integration of the EITC and the tax system, as under current law, makes sense.

We need say very little about the benefits of integration of the EITC with the tax system. As noted, measurements for income transfer programs will be closely related to measurements for the tax system, and coordination among income transfer and tax programs is important.

We also need say little about the problem of the use of different measurement criteria. Alstott argues that, ideally, a separate transfer program might use different (more accurate) measurement criteria than an integrated program would. The arguments on this issue are essentially the same as they were in the case of food stamps. Finally, discussion of the tradeoff between participation and compliance also tracks the arguments made with respect to food stamps.

The key difference between the EITC and the FSP, in our view, is the need for responsiveness in each program. We argued that if responsiveness is important for the FSP, integrating that program into the tax system would mean a sufficient sacrifice in responsiveness that integration would not be desirable. We believe that responsiveness is not as important for the EITC and, therefore, integration of the EITC into the tax system makes sense.

The core reason we believe this to be true is that the EITC is a wage supplement. Individuals are entitled to the EITC based on their wage income. It provides a supplement only to the working poor; nonworking individuals are not eligible. This is in contrast to food stamps, where eligibility does not depend on employment. The EITC also deviates in this respect from TANF and other transfer programs, for which individuals who do not work remain eligible (subject to incentives in the system to obtain work).

As a wage supplement, the EITC is not designed to be, and indeed cannot be, responsive to individuals in dire need. It cannot help individuals who have lost their jobs, who cannot find work, or otherwise need help immediately. Moreover, as a wage supplement, it is not designed to be temporary. Low-wage workers may remain that way for long periods of time, unlike (one hopes) those individuals who need emergency assistance. To this end, the EITC has no time cutoff, unlike TANF.

There is also evidence that the EITC functions well as an annual payment. Several studies indicate that individuals use the EITC as a form of forced savings and tend to invest the lump sum payments in ways that they would have been unlikely to invest monthly payments. For example, Jennifer Romich and Thomas Weisner find that families are more likely to purchase durable goods and make large purchases with the EITC than with other funds. Other studies have found that the lump sum EITC gives recipients an opportunity to make desirable changes in economic behavior.

As with food stamps, we would like to have more information before drawing conclusions about the need for responsiveness in the program. For example, we have been unable to find any information about EITC participation spells comparable to the evidence on food stamp spells. We also cannot conclude from the fact that the EITC is nonresponsive now that it would not be more

successful if it were responsive. What we observe may be an artifact of the fact that the program is integrated into the tax system, and may have little to do with its ideal structure.

More importantly, the whole notion of responsiveness of a given program is dependent on other programs. For example, the EITC might be able to be a nonresponsive wage subsidy only because other programs like food stamps and TANF are responsive. The arguments we have made are based on the existing programs rather than on a theory of which programs or how many programs should be responsive to short-term needs. Our conclusion thus remains, to a great extent, contingent on the broader institutional framework of the society within which the EITC operates. Nevertheless, given the basic framework of existing programs, integration of the EITC makes sense. Other programs can act to reduce emergency needs, allowing the EITC to be structured as a less responsive, but more efficient, wage subsidy.

D. Summary and Comments on the Relationship to Overall Reform

The arguments on integrating the FSP and the EITC into the tax system depend on specialization and coordination. There are strong reasons why integration makes sense given the specialization of the tax system (in measuring and processing relevant attributes and disbursing funds) and the benefits of coordination. A first approximation of the accuracy and complexity of the FSP and EITC supports this argument. Nevertheless, some elements of transfer programs could not easily be implemented by the IRS because it lacks the required level of expertise—the cost savings from tax system implementation would not be worth the reduced accuracy. In particular, the tax system is not designed to be able to respond quickly to changes in well-being, and some sort of transfer system is likely necessary to do this. To the extent food stamps are a necessary component of a quick response system, integration is not desirable.

More generally, we have tried to show how the integration decision depends on the tradeoff between coordination and specialization. Integrating most or all welfare programs into the tax system, such as through a negative income tax, is attractive because the tax and transfer systems rely on similar variables. A single agency can specialize in those variables and coordinate among the various programs. For example, one can imagine a broad system of refundable tax credits (or a negative income tax of another sort) combined with an emergency welfare system designed to reduce short-term needs. Such a system might be able to take advantage of the benefits of tax-based delivery of funds combined with responsiveness to short-term needs.

There may be problems with integration, however, because many elements of the transfer system could not be well-replicated within the tax system without a significant loss of specialization. For food stamps, the problems lie in the provision of in-kind benefits and rapid response. Study of other elements of the transfer system may reveal similar problems. One problem we can foresee is that the tax system may not have the expertise to exercise significant discretion regarding many problems of the transfer system. That is, a program that requires field agents to make non-income-based decisions about eligibility may not be ideally suited for integration. For example, a program might provide housing benefits for families on the condition that the children make good faith efforts to attend school. If the decision on whether the children have met this requirement involves discretion, integration into the tax system would be a bad idea.

Although the balance in any particular area would depend on the specific program needs and circumstances, integration may be worthwhile notwithstanding these problems. The benefits of specialized measurement and coordination may be greater than the costs. As noted in the comparison of the costs of administering food stamps and the EITC, there may be enormous savings from integration.

Conclusion

Rather than summarize what has already been said, let us suggest further avenues for research. First, as noted, we would like to extend the institutional framework to include agency costs and public choice concerns. The departmentalization framework gives us some valuable intuitions but many of these intuitions may change once agency costs and public choice concerns are incorporated.

We suspect, for example, that agency costs might explain some of the structure of government better than team theories of organizations. In particular, tax expenditures tend to be redundant in the sense that they grant to the IRS authority to implement a program that is within the expertise of another agency and often has close substitutes at the other agency. Team theory rationales cannot easily make sense of this pattern. An agency cost analysis, however, might. The idea would be that having multiple agents perform similar tasks allows Congress or the President to monitor the agents better and to have the ability to threaten each one with removal of the program. This might improve performance notwithstanding the increase in specialization and coordination costs. Further research is needed in this area.

Second, we would like to see more examples analyzed. This Article has used the EITC and food stamps as its primary examples. Transfer systems, however, provide a natural case for integration, and analysis of other programs may be more difficult. For example, there are numerous education, healthcare, and housing subsidies in the current law, and analysis of these provisions would be useful both for understanding the merits of the particular provisions and also for developing intuitions about the tax expenditure problem more generally. In addition, we would like to understand better the mechanisms used to coordinate policy in these areas and also in the private sector.

Third, nothing in this Article's framework limits its application to taxation. The same analysis might be used, for example, to determine whether antitrust policy is best implemented in the Justice Department or the Federal Trade Commission, or both. Similarly, this mode of analysis might help to determine whether environmental policy should be implemented through the Environmental Protection Agency, the Department of the Interior, the Army Corps of Engineers, other agencies, or some combination thereof. Understanding how the analysis applies in these other settings might help in understanding the tax setting, but might also yield interesting insights in those areas themselves.

Finally, the purpose of this paper was to understand the proper scope of the "tax system." The analysis can go in the other direction as well. Perhaps some tax collection functions are best carried out by other agencies. For example, an agency with specialization in farmers might best collect

taxes that are particular to farmers. Many countries have more than one tax collection agency, and the analysis presented here might be helpful in explaining this.

More generally, however, the purpose of the framework presented in this Article is to prompt us to question—to reconsider the costs and benefits of our present institutions and to evaluate whether our present allocations of tasks and functions across government agencies make sense. One would hope that in many cases they do; our analysis of the EITC and the FSP suggest such an outcome. But it is also possible that, in some cases, they do not. To the extent that our framework can help to identify successful institutional structures and ferret out examples of misallocated institutional functions, it can help the government to make better use of the public fisc.

Republished by permission of The Yale Law Journal Company and William S. Hein Company from The Yale Law Journal, Vol. 113, pages 955-1028.