המוסד לביטוח לאומי
האגף למחקר ותכנון

העלמת ניכויי מס במקור
והתחמקות ממס שלא - נוכה

מאת:
גדעון יניב

ירושלים, ניסן תשמ"ח, מרץ 1988

מחקר מס' 42

המוסד לביטוח לאומי
האגף למחקר ותכנון

העלמת נכויי מס במקור
והתחמקות ממס שלא - נוכה

מאת:
גדעון יניב

ירושלים, ניסן תשמ"ח, מרץ 1988

מחקר מס' 42

מלבד Marrelli (1984) אשר בדק מספר היבטים של העלמת מס קנייה, הספרות הכלכלית העוסקת במרמת מס, שהחלה להתפתח במהירות בעקבות מאמרם החלוצי של Allingham ו-Sandmo (1972), מתרכזת בעיקר בהעלמת הכנסות של פרטים ובהשלכותיה על החלטות היצע העבודה ומדיניות המיסוי האופטימאלית. כפי שצוין כבר על ידי Cowell (1985), רובה של ספרות זו (כמו, למשל Weiss (1976), Andersen (1977), Baldry (1979) או Pencavel (1979)), מתעלם מן העובדה שחלקה העיקרי (אם לא כל) חבות המס של עובד שכיר מנוכה בדרך כלל במקור על-פי תקנות מס הכנסה[1]. בעוד שמערכת נכויי המס במקור מגבילה במידה רבה את אפשרויות ההעלמה של עובדים שכירים (הגבלה המהווה את אחת הסיבות, אם כי לא העיקרית, לכינונה), היא עלולה ליצור תמריץ למעבידיהם להעביר לרשויות סכום קטן יותר מסכום המס (או סכום דמי הביטוח הלאומי) שנוכה. תמריץ כזה יטה יותר להופיע ככל שקטן יותר מספר השכירים החייב, על פי תקנות המס, להגיש דו"ח מס הכנסה[2]. ואולם, גם אם חובת הגשת הדו"ח היא כללית (ובהנחה שכל השכירים מצייתים לה), מעבידים עשויים עדיין להמנע מהעברת מלוא נכויי המס לרשויות אם ידוע, שמחמת משאבים מצומצמים, אין ביכולתן של רשויות המס להצליב במידה מלאה את הצהרותיהם כנגד אלה של השכירים.

מטרת המחקר הנוכחי היא לזהות את הגורמים המשפיעים על היקף מרמת המס (הלא מתואמת) של עובדים ומעבידים תחת מערכת של נכויי מס במקור, כמו גם לבדוק את השפעת המערכת על סכום המס המתחמק מן הרשויות. המאמר פותח בנתוח החלטת המעביד ביחס לשאלה האם ובאיזו מידה להמנע מהעברת מלוא נכויי המס (או נכויי הביטוח הלאומי) לרשויות על ידי הצהרה מכוונת על תשלומי שכר הנמוכים מאלה ששולמו בפועל. גורם מכריע המשפיע על החלטת המעביד הוא טיב הקשר שבין פעילותו הלא-חוקית ובין העונש הצפוי לו אם פעילותו תתגלה על-ידי הרשויות. העונש עשוי להקבע על פי גודל <u>הצהרת החסר</u> על תשלומי השכר או על-פי <u>סכום המס</u> שלא הועבר, ובמקרה שבו המעביד חייב במס רווחים, להתחשב (בדרכים שונות) בעובדה שהצהרת חסר על תשלומי שכר מתבטאת בתשלום יתר של מס רווחים - או

----

1. Beckman (1983) מתייחס למערכת הנכוי בארצות-הברית כאל "חוט השדרה של מערכת מס ההכנסה על פרטים", בציינו שב-1981 הסתכם סך הנכויים במקור ב-261 מליון דולר מתוך חבות מס כוללת של 291 ביליון דולר.

2. בעוד שבארצות-הברית, לדוגמא, חייבים בהגשת דו"ח גם שכירים בעלי הכנסה נמוכה יחסית (יחידים שהכנסתם השנתית, מכל מקור שהוא, עלתה, ב-1986, על 3,560 דולר), הרי שבישראל חובת הגשת הדו"ח מוגבלת לשכירים בעלי הכנסות גבוהות במיוחד (יחידים שהכנסתם השנתית מעבודה עלתה ב-1986, על שווה הערך של 33,000 דולר) בתנאי שהם עובדים במשרה אחת בלבד ואינם זוכים להכנסה מאף מקור אחר.

להתעלם מעובדה זו. פרק II של המאמר מציג תיאור פורמאלי של בעיית המעביד, גוזר תנאים מספיקים לאי העברה מלאה של נכויי המס לממשלה, בוחן את השפעתם של תנאים אלה על ביקוש המעביד לעובדים ודן ביחסים האפשריים שבין היקף המרמה ובין שעורי הנכוי במקור, שעור מס הרווחים ורמת שכר העבודה תחת צורות ענישה אלטרנטיביות.

בהתייחסו להשפעה האפשרית על העלמת הכנסות של כינון מערכת נכויי מס במקור טוען Cowell (1985), שהמעלים המיומן יטול משרה נוספת (או יעבור למשרה אחרת) שההכנסה ממנה קשה אדמיניסטרטיבית למיסוי. בעוד שדרך פעולה זו נראית כאפשרות יחידה להעלמה כאשר תקנות מס הכהנסה מחייבות שכל חבות המס תנוכה במקור, הרי שאין היא בלעדית במצב שבו, כמו בארצות הברית, שעורי הנכוי במקור נמוכים משעורי המס הסופיים. אף כי בעל משרה יחידה לא יוכל גם אז להתחמק מתשלום מס על ידי הצהרת חסר (שכן דו"ח המס שלו חייב להיות מלווה באישור המעביד על הכנסותיו בפועל וסכום המס שנוכה), הרי שיהיה באפשרותו להתחמק מתשלום המס שלא נוכה - וזאת על ידי הימנעות מהגשת דו"ח לרשויות המס. פרק III של המאמר עוסק בקובעי התנהגות אפשרית זאת ובמיוחד בתלותה בשעורי הנכוי במקור - כמו גם בדרך שבה ההחלטה שלא להגיש דו"ח משפיעה על היצע העבודה של הפרט. לבסוף, פרק IV דן בשאלת יכולתה של מערכת הנכוי במקור להביא להקטנת סכום המס המועלם במשק. פרק V מסיים בסיכום התוצאות העיקריות ובמספר הערות רלבנטיות.

NATIONAL INSURANCE INSTITUTE
Bureau of Research and Planning

WITHHOLDING AND NON-WITHHELD TAX EVASION*

by
Gideon Yaniv

*Forthcoming in the Journal of Public Economics. I am grateful to two anonymous referees for valuable comments and suggestions on an earlier draft of this paper.

Jerusalem, March 1988 Discussion paper 42

I. INTRODUCTION

With the exception of Marrelli (1984), who examined some aspects of excise tax evasion, the literature on tax fraud, initiated by Allingham and Sandmo (1972), has been mainly concerned with individual income tax evasion and its implications on labor supply decisions and optimum taxation policy. As pointed out by Cowell (1985), most of this literature (e.g., Weiss (1976), Andersen (1977), Baldry (1979) or Pencavel (1979)) ignores the fact that the major part (if not all) of a wage-earner's tax liability is usually deducted at source by withholding regulations.[1] However, while greatly limiting wage-earners' possibilities of tax evasion (which is one, although not the main, reason for its establishment), a withholding tax system might provide incentives for withholding agents to remit to the government less than the amounts withheld. Such incentives are more likely to emerge the lower is the number of wage-earners obliged by tax regulations to file an income tax return.[2] Yet even if the obligation to file a return is of a universal nature (and supposing that all wage-earners comply with it), partial

[1]Peckman (1983) regards the withholding system in the United States as the "backbone of the individual income tax", stating that in 1981, withholdings brought in $261.0 billion out of a total tax liability of $291.1 billion.

[2]While in the United States, for example, tax returns must be filed even by relatively low-wage earners (single persons whose annual income, regardless of its source, exceeds $3,560 in 1986), in Israel, filing a return is obligatory only for extremely high-wage earners (single persons whose annual earnings exceed, in 1986, the equivalent of $33,000), providing that they hold a single job and do not receive income from any other source.

remittance of withheld taxes might still be practiced if limited resources constrain tax authorities from fully cross-checking individual returns against employers' declarations.

The purpose of the present paper is to inquire into the determinants and scope of the employer and employee (uncollaborated) tax fraud activity under a withholding tax system, as well as into the system's effect on the amount of tax escaping the tax collector. We begin by analyzing the employer's decision on whether and to what extent to avoid full remittance of his employees' withheld taxes by deliberately understating his actual wage payments. A crucial factor affecting the employer's decision is the form of the penalty that he should expect if his fraudulent behavior is detected by the authorities. The penalty may relate to the magnitude of his unstated wage payments or to the amount of non-remitted taxes, and, in the presence of profit taxation, takes account (in different ways) of the fact that understatement of wage payments results in overpayment of profit taxes - or disregard it. Section II of the paper provides a formal description of the employer's problem, derives conditions sufficient to induce less than full remittance of the amount withheld, examines their effect on the demand for labor and discusses the possible relationships between the magnitude of fraud, the withholding and profit tax rates and the competitive wage rate under alternative penalty schemes.

Considering wage-earners' reaction to the introduction of a withholding system, Cowell suggests that the skillful evader would take a second (or would switch entirely to another) job - the income from which is administratively difficult to tax. While this seems to be the sole opportunity for evasion when withholding regulations require that all taxes due will be deducted at source, it is not an exclusive one when, as is the case in the United States, the withholding tax rates are kept below the final rates. Although a single job holder would still be unable to evade taxes through underreporting (as his tax return must be accompanied by an employer's statement of actual earnings and the amount withheld), he might nevertheless evade the entire amount of his non-withheld taxes by avoiding filing a return altogether. Section III investigates the determinants of such behavior - in particular its dependence upon the withholding tax rate - and how deciding against filing a return affects the supply of labor. Finally, Section IV addresses the question of whther a withholding system actually helps to reduce the amount evaded. Section V concludes with a summary of main results and some related remarks.

II. EVASION OF TAX WITHHOLDINGS

Consider a competitive employer facing a fixed wage rate, w, per hour of employed labor, N, who is required by tax regulations to withhold a given proportion, t, of his total wage payments. Suppose that the employer considers the possibility of abusing the withholding system by remitting to the tax authorities, via understating his actual wage payments, less than the amount withheld. Cheating the authorities would expose him, of course, to the risk of being detected and punished. Denoting declared wage payments by $Z(<wN)$, the employer's net profits if not detected, π^{nd}, will be

$$\pi^{nd} = (1-\theta)V(N) - (1-t)wN - (t-\theta)Z \tag{1}$$

where θ represents a constant profit tax rate[3] and V(N) indicates the value of output, assumed to be accurately declared, and produced at decreasing marginal rates $[V'(N)>0$ and $V''(N)<0]$ by labor input only.

If, on the other hand, his fraudulent behavior is detected, the employer will have to pay a penalty, F, which we will assume to be linearly related either to the magnitude of unstated wage payments or to the amount of non-remitted tax with-

---

[3]In the absence of profit taxation ($\theta=0$), π^{nd} reduces into $V(N)-(1-t)wN-tZ$. The implications of this special case are derived later in this section as a private case of the more general formulation.

holdings.[4] These alternative forms of punishment are analogous to those proposed, respectively, by Allingham and Sandmo (1972) and Yitzhaki (1974) with respect to individual income tax evasion. However, since understatement of wage payments results in overpayments of profit taxes, we will first assume that if the tax authorities get to know the exact amount of his wage payments, the employer will be refunded for his profit tax overpayments of $\theta(wN-Z)$,[5] or will actually be penalized in accordance with

$$F = (T-\theta)(wN-Z) \tag{2}$$

where $T>t$ (and, of course, $T>\theta$). Under the former penalty scheme, T is a constant ($=\bar{T}$), independent of variations in t, while under the latter, $T=\lambda t$, where $\lambda>1$. Subtracting now (2) from (1), the employer's net profits if detected, π^d, will be

$$\pi^d = (1-\theta)[V(N)-wN] - (T-t)(wN-Z) \tag{3}$$

Suppose further that the employer is risk-averse, that his utility function, U, has net profits as its only argument and that he chooses Z* and N* so as to maximize the expected utility of his prospect

$$EU(\pi) = (1-p)U(\pi^{nd}) + pU(\pi^d) \tag{4}$$

---

[4]Under both the U.S. and the Israeli tax laws, an employer who willfully (but not necessarily with fraudulent purpose) fails to remit his employees' tax withholdings must pay back twice the amount of the tax evaded. If, however, his failure is due to fraudulent intentions, he must, under the Israeli tax law, pay back twice the amount of the unstated wage payments, and under both laws may also be subject to a criminal penalty carrying fixed heavy fines and/or imprisonment.

[5]Neither the U.S. nor the Israeli tax law refers to this issue specifically. Both laws specify, however, that any amount overpayed should be returned upon tax reconciliation. The interesting question, of course, is whether this applies to the case where overpayment is due to an attempt to defraud the system elsewhere. Accountants who have been consulted with regard to this issue believe that it does. In what follows we also examine the possibility that the penalty is assessed on the <u>net</u> amount evaded or that profit tax overpayments are disregarded by the tax authorities.

where p denotes the (exogenously given) probability of being detected. The first-order conditions for an interior maximum will then be

$$\frac{d[EU(\pi)]}{dZ} = -(t-\theta)(1-p)U'(\pi^{nd}) + (T-t)pU'(\pi^{d}) = 0 \tag{5}$$

$$\frac{d[EU(\pi)]}{dN} = [(1-\theta)V'(N)-(1-t)w](1-p)U'(\pi^{nd}) + \\ + \{(1-\theta)[V'(N)-w]-(T-t)w\}pU'(\pi^{d}) \tag{6}$$

A first glance at equation (5) reveals that a necessary prerequisite for evading withheld taxes is that $t>\theta$, since underreporting actual wage payments increases profit tax liabilities.[6] The sufficient condition for adopting this practice is, however, stricter, as it requires that $t > pT+(1-p)\theta$.[7] That is, the marginal gain per underreported dollar actually paid to labor should exceed the expected marginal loss per that dollar in order for evasion incentives to arise.

Substituting (5) into (6) reveals now that at the optimum

$$V'(N) - w = 0 \tag{7}$$

which is the usual marginal condition for profit maximizing employment of labor input. We thus conclude that the optimal employment level is independent of fraud behavior (as well as of the withholding tax rate), as long as the latter is opti-

[6]The likelihood of this condition to hold increases greatly by the fact that the effective tax rates on profits often lie drastically below the statutory rates due to generous tax credits and depreciation allowances. Peckman points out that while the statutory rate in the United States amounted in 1982 to 46 percent, the effective rate was, on average, as low as 13.1 percent (an important characteristic of the recent tax reform in the United States is the elimination of many of these tax shelters, compensated by a reduction in the top tax rate to 34 percent). The withholding rates, on the other hand, are usually not subjected to most of the deductions allowed to many employees (particularly those who own homes), which are itemized upon filing a final tax return.

[7]To see this, notice that $Z^*<wN$ implies that at the optimum $\pi^{nd}>\pi^{d}$, hence, by risk-aversion, that $U'(\pi^{nd})<U'(\pi^{d})$, which requires, by (5), that $(t-\theta)(1-p)>(T-t)p$.

mal.[8] The inverse, however, does not hold.

Turning now to examine the employer's reaction to changes in his environment, it can easily be verified that an increase in the law enforcement parameters (the probability of detection or the penalty multipliers) would discourage tax fraud, as is intuitively expected (and almost always obtained) in tax evasion models concerned with risk-averse taxpayers. The interesting effects to focus attention on are those produced by variations in the withholding and profit tax rates and the competitive wage rate, the latter of which would also affect evasion indirectly through its effect on optimal employment. Considering first the employer's response to a change in the withholding tax rate, we obtain, using Cramer's Rule, that

$$\frac{dZ^*}{dt} = \frac{EU'(\pi)}{\Delta} + \frac{(t-\theta)(wN-Z)(1-p)U'(\pi^{nd})}{\Delta}\left[R_A(\pi^d)-R_A(\pi^{nd})\right] \qquad (8)$$

when $T=\bar{T}$, and that

$$\frac{dZ^*}{dt} = \frac{EU'(\pi)-\lambda pU'(\pi^d)}{\Delta} + \frac{wN-Z}{\Delta}\left[(t-\theta)(1-p)U''(\pi^{nd})+t(\lambda-1)^2pU''(\pi^d)\right] \qquad (9)$$

when $T=\lambda t$, where $R_A(\pi)=-U''(\pi)/U'(\pi)>0$ is the Arrow-Pratt absolute risk-aversion measure, and $\Delta \equiv (t-\theta)^2(1-p)U''(\pi^{nd}) + (T-t)^2U''(\pi^d) < 0$ is a second-order condition for the maximization of expected utility.

The right-hand-side of equations (8) and (9) consists of two terms, which may be regarded as substitution and income effects, respectively. The former arises due to the fact that a change in the withholding tax rate affects the price of a re-

[8]This result (which is easily verified by substituting (7) into (6) to obtain (5) - and vice versa) is due to the fact that, with the exception of its effect on legitimate net profits, $(1-\theta)[V(N)-wN]$, a change in N affects expected utility in exactly the opposite way as a change in Z. Since the latter effects must sum to zero at the optimum - condition (7) arises. This result is analogous to Marrelli's (1984) finding that profit-maximizing production is independent of optimal excise tax evasion.

ported relative to an unreported dollar of wage payments, $(t-\theta)/p(T-\theta)$.[9] The latter stems from the fact that a change in the withholding tax rate affects net profits at both states of the world (detection and non-detection) for any given level of initial declaration. Under the accepted assumption of decreasing absolute risk-aversion $[R_A(\pi^d)>R_A(\pi^{nd})]$, the sign of (8) is unambiguously negative, as the income effect operates in the same direction as the substitution effect. In other words, an increase in the withholding tax rate would not only increase the price of a reported relative to an unreported dollar (or the profitability of under-reporting on the margin), but would also increase profits at both states of the world, thus tending, by the customary interpretation of the income effect, to stimulate risk-taking.[10]

The sign of (9), however, is indeterminate. While the substitution effect is still negative,[11] the income effect is unambiguously positive - regardless of risk-aversion behavior and in spite of the fact that profits at the alternative states of the world vary in opposite directions. Obviously, the customary interpretation of the income effect is inapplicable to this case, nor will it apply to most of the other comparative statics results derived hereafter.[12] Nevertheless, all in-

[9]Alternatively, the substitution effect may be viewed as resulting from a change in the marginal return per unreported dollar, t, relative to the expected marginal cost per this dollar, $pT+(1-p)\theta$.

[10]As first argued by Allingham and Sandmo (1972), an increase in the regular tax rate - formally found to have a positive income effect on a taxpayer's declaration - would make the taxpayer less wealthy (whether being detected or not), which, under decreasing absolute risk-aversion, tends to reduce evasion. This interpretation of the income effect has since been adopted by other writers addressing the issue of tax evasion, e.g., Andersen (1979), Pencavel (1979) and Koskela (1983). Notice also that contrary to (8), the total effect of an increase in the regular tax rate on a taxpayer's declaration is ambiguous.

[11]This is so since (5) implies that $EU'(\pi)-\lambda U'(\pi^d) = (1-p)U'(\pi^{nd})\theta/t > 0$.

[12]While Allingham and Sandmo's interpretation of the income effect seems to hold for parameter variations producing equal income changes in both states of the world,

come effects derived in this section can still be rationalized consistently if one perceives risk-aversion as a stimulant for cushioning (or for dispersing over both states) – via an adjustment in declaration - any exogenous shock affecting either state of the world profits.[13] Under a "shock cushioning" perception, a parameter change which affects both states of the world profits would give rise to two distinct income effects on declaration. These effects would act in the same direction - regardless of risk-aversion behavior - if the parameter change affects the states of the world profits in opposite directions. Thus, an increase in the withholding tax rate, which increases profits in the state of non-detection but - due to the fact that the penalty rate is proportional to the former - decreases profits in the state of detection, would unambiguously encourage declaration (equation (9)). However, a parameter change which affects both states of the world profits in the same direction would give rise to opposing income effects - the magnitude of each depends (directly) on the size of the corresponding profit change as well as on the extent of risk-aversion. Therefore, an equal size increase in profits resulting from a rise in the withholding tax rate when the penalty rate is independent of the former, would - under decreasing absolute risk-aversion - discourage declaration (equation (8)).

The ambiguous relation between wage payments declaration and the withholding tax rate prevailing under a penalty scheme which relates to the magnitude of non-remitted taxes will, however, be eliminated if the penalty is assessed on the net

---

it is inapplicable to most of the cases involving unequal income changes, where decreasing absolute risk-aversion is either irrelevant to the determination of the income effect or produces a result which is inconsistent with their interpretation. A formal proof of this argument is provided in a recent note by myself (Yaniv, 1988).

[13]Thus, a parameter change which increases (decreases) profits in the state of non-detection or decreases (increases) profits in the state of detection would encourage (discourage) declaration, as this would help to decrease (increase) profits in the former state and to increase (decrease) profits in the latter.

amount evaded. That is, if

$$F = \lambda(t-\theta)(wN-Z) \tag{2'}$$

which implies, when compared with (2), that if detected, the employer is more than refunded for his profit tax overpayments.[14] The employer's net profits in case of detection will thus be

$$\pi^d = (1-\theta)V(N) - [1-t+\lambda(t-\theta)]wN + (\lambda-1)(t-\theta)Z \tag{3'}$$

generating

$$\frac{d[EU(\pi)]}{dZ} = (t-\theta)[-(1-p)U'(\pi^{nd}) + (\lambda-1)pU'(\pi^d)] = 0 \tag{5'}$$

as a first-order condition (with respect to Z) for the maximization of expected utility, and $p\lambda<1$ as a sufficient condition for fraudulent behavior. Applying Cramer's Rule to (5') we obtain

$$\frac{dZ^*}{dt} = \frac{wN-Z}{t-\theta} \tag{9'}$$

which is unambiguously positive due to an income effect only.[15] The negative substitution effect prevailing under penalty function (2) has been eliminated, since the relative price of a reported and an unreported dollar, $(t-\theta)/p\lambda(t-\theta)$, is now independent of t.[16]

Considering now the employer's response to an increase in the profit tax rate, we obtain, under penalty function (2)

---

[14]Notice that when (2) relates to evaded taxes (i.e., when $T=\lambda t$), it becomes $F = (\lambda t-\theta)(wN-Z)$, which is not identical to (2').

[15]The income effect term is considerably simplified in this case since the joint derivative of $EU(\pi)$ with respect to Z and t is now linearly related to the second-order derivative of $EU(\pi)$ with respect to Z.

[16]It is interesting to note, in analogy, that Yitzhaki (1974) reveals a positive relation between a taxpayer's income declaration and the regular tax rate - when the penalty rate varies with the latter. His result, however, is contingent upon decreasing or constant absolute risk-aversion.

$$\frac{dZ^*}{d\theta} = -\frac{(1-p)U'(\pi^{nd})}{\Delta} \{1 + (t-\theta)[(V(N)-wN)R_A(\pi^d) - (V(N)-Z)R_A(\pi^{nd})]\} \qquad (10)$$

which reflect a positive substitution effect - due to a reduction in the price of a reported relative to an unreported dollar - and an indeterminate income effect, in spite of the fact that an increase in θ reduces both states of the world net profits. The intuitive explanation of this result, in accordance with our current interpretation, is that an increase in θ generates actually two opposing income effects: while decreasing absolute risk-aversion acts to encourage declaration, the greater reduction in net profits in the state of non-detection acts to discourage it. Consequently, the sign of (10) is ambiguous.

This ambiguity is somewhat relaxed under penalty function (2'), which yields[17]

$$\frac{dZ^*}{d\theta} = -\frac{(t-\theta)(1-p)U'(\pi^{nd})}{\Delta} \{[V(N)-Z-\lambda(wN-Z)]R_A(\pi^d) - [V(N)-Z]R_A(\pi^{nd})\} \qquad (10')$$

for which the substitution effect is eliminated, as the relative price of a reported and an unreported dollar becomes independent of θ. The (net) income effect is still ambiguous for $\lambda<\lambda^*=[V(N)-Z]/(wN-Z)$, i.e., as long as the penalty multiplier is less than the ratio between declared profits and undeclared wages, but negative - regardless of risk-aversion behavior - otherwise. The latter, which is likely to occur at relatively high levels of evasion and employment, is due to the fact that an increase in θ, while reducing net profits in the state of non-detection, does not affect (when $\lambda=\lambda^*$) or increases (when $\lambda>\lambda^*$) net profits in the state of detection, thus producing only one or two same direction income effects, respectively. Consequently, the sign of (10') would be negative.

Although being less plausible, we may still wish to examine the case where a detected employer is not credited in any way for his profit tax overpayments, i.e.,

---

[17]Hereafter Δ is used to indicate, in general, the second-order derivative of $EU(\pi)$ with respect to Z. Although its value varies slightly with variations in the penalty function - its negative sign is always preserved.

where

$$F = T(wN-Z) \tag{2"}$$

The employer's net profits if detected will then be

$$\pi^{d} = (1-\theta)V(N) - (1+T-t)wN + (\theta+T-t)Z \tag{3"}$$

yielding as a first-order condition (with respect to Z)

$$\frac{d[EU(\pi)]}{dZ} = -(t-\theta)(1-p)U'(\pi^{nd}) + (\theta+T-t)pU'(\pi^{d}) = 0 \tag{5"}$$

which allows for tax evasion if $t>pT+\theta$. The effect of an increase in the profit tax rate is now given by

$$\frac{dZ^*}{d\theta} = -\frac{1}{\Delta}\{EU'(\pi) + (t-\theta)[V(N)-Z](1-p)U'(\pi^{nd})[R_A(\pi^{d})-R_A(\pi^{nd})]\} \tag{10"}$$

for which the substitution and income effects (given decreasing absolute risk-aversion) are both positive. The former because of a reduction in the relative price, $(t-\theta)/pT$, whereas the latter due to an equal fall in net profits in both states of the world. However, while totally eliminating the ambiguity from the relation between declared wage payments and the profit tax rate, the present case preserves the negative substitution effect of a change in t when $T=\lambda t$, thus retaining an ambiguous relation between declared wage payments and the withholding tax rate. Notice, on the other hand, that penalty function (2") applies also to the case where profits are not subjected to taxation at all ($\theta=0$), for which the substitution effect of a change in t disappears again, given that $T=\lambda t$. The negative relation between declared wage payments and the withholding tax rate identified for $T=\bar{T}$, prevails, however, under both cases captured by penalty function (2").

Table 1 summarizes the possible effects on declared wage payments of variations in the withholding and profit tax rates under the alternative penalty schemes discussed in this section. As regards the effect on declaration of a change in

Table 1: Effects on Declared Wage Payments of Variations in the Withholding and Profit Tax Rates

penalty function \ parameter change		t			θ		
		substitution	income	total	substitution	income	total
$F=(\bar{T}-\theta)(wN-Z)$		<0	<0	<0	>0	$\gtreqless 0$	$\gtreqless 0$
$F=(\lambda t-\theta)(wN-Z)$		<0	>0	$\gtreqless 0$	>0	$\gtreqless 0$	$\gtreqless 0$
$F=\lambda(t-\theta)(wN-Z)$		$=0$	>0	>0	$=0$	<0 $\gtreqless 0$	if $\lambda \geq \lambda^*$ if $\lambda<\lambda^*$
$F=\bar{T}(wN-Z)$	$\theta=0$	<0	<0	<0	-	-	-
	$\theta>0$	<0	<0	<0	>0	>0	>0
$F=\lambda t(wN-Z)$	$\theta=0$	$=0$	>0	>0	-	-	-
	$\theta>0$	<0	>0	$\gtreqless 0$	>0	>0	>0

the competitive wage rate, we obtain, taking account of its effect on employment

$$\frac{dZ^*}{dw} = \frac{(t-\theta)N(1-p)U'(\pi^{nd})}{\Delta}\{[1-\theta+(T-t)(1-\varepsilon)]R_A(\pi^d) - [1-t+(t-\theta)\varepsilon]R_A(\pi^{nd})\} \quad (11)$$

given penalty function (2), where $\varepsilon=-(dN/dw)(w/N)>0$ is the elasticity of labor demand with respect to the wage rate. Clearly, an increase in w would not generate any substitution effect, since the relative price of a reported and an unreported dollar is independent of the wage rate or the level of employment. While always reducing profits in the state of non-detection, an increase in w would reduce profits in the state of detection only if $\varepsilon<\varepsilon^*=1+(1-\theta)/(T-t)$, increasing or not affecting profits if $\varepsilon\geq\varepsilon^*$. Moreover, the reduction in profits in the latter state (when $\varepsilon<\varepsilon^*$) would be greater than or equal to the reduction in profits in the former state if $\varepsilon\leq 1$, and smaller than the former if $1<\varepsilon<\varepsilon^*$. Consequently, the sign of (11) would be positive or ambiguous, given decreasing absolute risk-aversion, when $\varepsilon\leq 1$ or $1<\varepsilon<\varepsilon^*$, respectively, and negative, regardless of risk-aversion behavior, when $\varepsilon\geq\varepsilon^*$. This conclusion holds for penalty functions (2') and (2") as well, only the value of ε^* would change to $[1-t+\lambda(t-\theta)]/(\lambda-1)(t-\theta)$ and $(1+T-t)/(\theta+T-t)$, respectively.

Finally, it is interesting to inquire whether the results obtained so far still hold upon relaxing the linearity of the withholding tax or the penalty schedule - and in particular, upon replacing it by a progressive one. Since a progressive tax schedule applies to individual earnings (and not to wage payments as a whole), a slight modification of the employer's problem is needed. Suppose then that the employer of M identical workers is required by tax regulations to withhold the amount of tW^{α} from each of his workers earnings, where W denotes a competitive wage rate per worker, and $\alpha>1$ (but $\alpha tW^{\alpha-1}<1$). Evasion of tax withholdings may now take the form of either underreporting the actual wage rate or the actual number of workers. However, since each worker must be provided with a statement of his actual wages and the amount of tax withheld (and since the tax authorities might be able to detect the competitive wage rate with negligible costs), a dishonest (but rational) employer would apparently adhere to the latter option, reporting the employment of $G(\leq M)$ workers only.

Suppose now that the penalty schedule is linear or progressive with respect to either the number of unreported workers or the amount of non-remitted taxes, and that profit tax overpayments are refunded upon detection. That is, suppose that

$$F = T(M-G)^{\beta} - \theta W(M-G) \quad (2''')$$

where $T > tW^{\alpha}/\beta(M-G)^{\beta-1}$,[18] and $\beta \geq 1$. Under the former schedule $T = \bar{T}$, while under the latter $T = \lambda(tW^{\alpha})^{\beta}$, where $\lambda > 1/\beta[tW^{\alpha}(M-G)]^{\beta-1}$. Alternative net profits will thus be

$$\pi^{nd} = (1-\theta)V(M) - W(M-\theta G) + tW^{\alpha}(M-G) \quad (1''')$$

$$\pi^{d} = (1-\theta)[V(M)-WM] + tW^{\alpha}(M-G) - T(M-G)^{\beta} \quad (3''')$$

implying that the first-order conditions for the maximization of expected utility are

$$\frac{d[EU(\pi)]}{dG} = -(tW^{\alpha}-\theta W)(1-p)U'(\pi^{nd}) + [\beta T(M-G)^{\beta-1}-tW^{\alpha}]pU'(\pi^{d}) = 0 \quad (5''')$$

$$\frac{d[EU(\pi)]}{dM} = [(1-\theta)V'(M)-W+tW^{\alpha}](1-p)U'(\pi^{nd}) + $$
$$+ \{(1-\theta)[V'(M)-W]+tW^{\alpha}-\beta T(M-G)^{\beta-1}\}pU'(\pi^{d}) = 0 \quad (6''')$$

It can easily be verified now that relaxing the linearity of the tax and penalty schedules would not abolish the independency of the employment decision (substituting $V'(M)=W$ into (6''') yields (5''')). Also, given that $\beta=1$ (i.e., that the penalty function is linear), a non-linear tax schedule would still preserve the qualitative implications on declaration of variations in the (withholding and

---

[18]This is so as the marginal penalty per unreported worker, $\beta T(M-G)^{\beta-1}$, must exceed the marginal return, tW^{α}.

profit) tax and penalty parameters.[19] Thus, for example, an increase in t, when $T=\bar{T}$, would still raise both states of the world profits by the same amount; as well as increase the price of a reported relative to an unreported worker, $(tW^{\alpha}-\theta W)/p(\bar{T}-\theta W)$, producing unambiguous negative income and substitution effects on optimal declaration. However, the introduction of a non-linear tax schedule would obscure the employer's response to a change in the competitive wage rate: not only would an increase in W produce now a (negative) substitution effect on declaration, it would also have an indistinctive effect on net profits in case of detection.

The above conclusions would also hold under the alternative assumptions regarding the treatment of profit taxes - only the substitution effect of a change in W would be eliminated whenever the substitution effect of a change in t would. Given that $\beta>1$, these conclusions would continue to hold as long as the penalty scheme relates to underreporting. However, if the employer is penalized with respect to non-remittance, the substitution effect of a change in t or W would no longer be eliminated. Instead, when the penalty scheme relates to the net amount evaded (or in the absence of profit taxation), the substitution effect would become positive, since an increase in t or W would <u>reduce</u> the price of a reported relative to an unreported worker, $(tW^{\alpha}-\theta W)/p\lambda(tW^{\alpha}-\theta W)^{\beta}(M-G)^{\beta-1}$. On the other hand, an increase in t or W would produce an indeterminate substitution effect on declaration when profit tax overpayments are simply refunded or disregarded upon detection. While the positive substitution effect emerging in the former cases would help to strengthen the positive relation between declaration and the withholding rate revealed before, the indeterminate substitution effect characterizing the latter cases would only aggravate an already existing ambiguity. We may thus concluse that, with the excep-

[19]This conclusion is analogous to Pencavel's (1979) finding that as long as a taxpayer's declaration is independent of his labor supply decision, a non-linear tax structure would not affect the sign implications of changes in the tax and penalty parameters derived under a linear tax structure.

tion of a change in the competitive wage rate, the introduction of non-linear tax and penalty schedules would not affect the overall qualitative implications of changes in the employer's environment prevailing under a linear tax and penalty system.

III. EVASION OF NON-WITHHELD TAXES

As pointed out by Peckman (1983), many wage-earners in the United States do not have their taxes fully withheld because the withholding rates do not reach as high as the final tax rates.[20] A wage-earner whose estimated tax exceeds the withheld tax (by more than a specified amount) is required to pay the difference by filing a declaration of estimated tax. Final reconciliation between the actual tax liability and prepayments is made upon filing the final tax return. Clearly, a withholding system does not provide a single job holder with the opportunity of evading taxes through underreporting of actual earnings.[21] He may nevertheless evade the entire amount of his non-withheld taxes by avoiding filing a declaration and a final return altogether. The determinants of this behavior are worth inquiring.

Consider a risk-averse wage-earner holding a single job which pays w per hour of work. Suppose that his employer withholds a proportion t of his earnings for tax purposes, yet he actually faces a higher tax rate, τ, which will be applied to his total earnings upon filing a final tax return.[22] Denoting hours of work by H,

---

[20]While the maximum tax rate on earned income - prior to the recent tax reform - was 50 percent, the maximum withholding rate was only 37 percent. In Israel, however, the withholding rates amount roughly to the final rates, exempting single job holders with no other income from the obligation of filing a return.

[21]A multiple job holder may still underreport his actual earnings by not reporting employment in one or more of the jobs he holds. This case is examined by Cowell (1985).

[22]We assume, of course, that τ represents an after credits final rate (i.e., after itemized deductions allowed on the final report have been taken account of).

the worker's net earnings after tax reconciliation, $\hat{I}$, will be

$$\hat{I} = (1-\tau)wH \tag{12}$$

Suppose, however, that the worker considers the possibility of not filing a return. If his avoidance is not detected, his net earnings, I^{nd}, will be

$$I^{nd} = (1-t)wH \tag{13}$$

while if his avoidance is detected, he will be obliged to pay a penalty, f, which we will assume to be proportional to the amount of evaded taxes.[23] That is,

$$f = \delta(\tau-t)wH \tag{14}$$

where $\delta>1$. The worker's net earnings if detected, I^d, will thus be

$$I^d = [1-t-\delta(\tau-t)]wH \tag{15}$$

Suppose now that the worker's utility function, ψ, is defined upon net earnings and leisure, and that his time constraint, T, and the probability of being detected, ρ, are exogenously fixed. Assuming, at first stage, that the volume of his employment is determined by the employer - at the level of $\bar{H}$ hours - the worker will decide against filing a return if

$$E\psi(I, T-\bar{H}) = (1-\rho)\psi(I^{nd}, T-\bar{H}) + \rho\psi(I^d, T-\bar{H}) > \psi(\hat{I}, T-\bar{H}) \tag{16}$$

which is clearly satisfied for $\rho=0$. In order to derive other implications of this condition we may approximate $\psi(I^{nd}, T-\bar{H})$ and $\psi(I^d, T-\bar{H})$ by second-order Taylor

[23]Since tax evasion is not due to underreporting, the penalty scheme can only relate to the former. Under the U.S. tax law, there is a combined penalty of 5 percent of the unpaid tax for each month the return was late (for no reasonable cause), consisting of 4.5 percent for filing late plus a 0.5 percent for paying late. When the maximum penalty for filing late (22.5 percent) has been reached, the penalty for paying late continues until it too reaches its maximum (25 percent). Therefore, when the maximum combined penalty is assessed, the total rate will be 47.5 percent of the tax due. Willful failure to file a return may also be subject to a criminal penalty carrying a fine of not more than $25,000, or imprisonment for not more than one year, or both.

expansions around $\psi(\hat{I}, T-A)$, obtaining, after rearranging terms[24]

$$E\psi(I, T-A) - \psi(\hat{I}, T-A) \cong (\tau-t)wA\{(1-\rho\delta)\psi_I(\hat{I}, T-A) + + \frac{1}{2}[1-\rho\delta(2-\delta)](\tau-t)wA\psi_{II}(\hat{I}, T-A)\} \quad (17)$$

Substituting (17) into (16), the nonfiling condition becomes

$$(\tau-t)wAR_A(\hat{I}, T-A) < \phi(\rho, \delta) \quad (18)$$

where $\phi(\rho, \delta) = 2(1-\rho\delta)/[1-\rho\delta(2-\delta)]$. We first notice that a necessary prerequisite for nonfiling is that $\phi(\rho, \delta)>0$, which requires - since the denominator of $\phi(\rho, \delta)$ is positive[25] - that $\rho\delta<1$.[26] Secondly, since $\phi_\rho<0$ and $\phi_\delta<0$, it follows, as intuitively expected, that the incentive to file is lower the lower are the law enforcement parameters ρ and δ. Thirdly, and also as expected, the incentive to file is lower the lower is risk-aversion when complying with tax regulations. However, contrary to intuition, the incentive to file will be lower the closer is the withholding tax rate to the actual rate faced by the taxpayer. The possible explanation for this implication is that the reduction in the cost of filing (taxes payed beyond the amount withheld) associated with a higher withholding rate is more than offset by the reduction in the cost of nonfiling (expected penalty). Also, given decreasing absolute risk-aversion, the incentive to file will be lower the lower is the actual tax rate, as the latter will not only help to reduce the cost of nonfiling but will also increase legitimate net earnings and thus risk-taking. Finally, the incentive to file will be lower the lower or the higher is the wage rate - and thus gross earnings - if the elasticity (in absolute terms) of risk-aversion with

---

[24]Notice that $I^{nd}-\hat{I} = (\tau-t)wA$ and $I^d-\hat{I} = -(\delta-1)(\tau-t)wA$.

[25]Subtracting from and adding ρ to $1-\rho\delta(2-\delta)$ yields $1-\rho+\rho(1-\delta)^2 > 0$.

[26]Alternatively, notice that $\hat{I}$ can be expressed as $[1-t-(\tau-t)]wA$, while expected net earnings when not filing, I, is $[1-t-\rho\delta(\tau-t)]wA$. Hence, when $\rho\delta=1$, the nonfiling option is "actuarilly fair", and should be rejected by a risk-averse taxpayer. A necessary, yet not sufficient, condition for non-rejection is thus $\rho\delta<1$.

respect to legitimate net earnings is smaller or greater than unity, respectively. This is so since a lower (higher) level of gross earnings affects the incentive to file in opposite directions, decreasing (increasing) the cost of nonfiling on the one hand, but discouraging (encouraging) risk-taking on the other. These conclusions would also hold in the presence of a progressive tax schedule under which the employer is required to withhold the amount of $t(w\bar{H})^{\alpha}$ while the worker must add $(\tau-t)(w\bar{H})^{\alpha}$ upon filing a tax return - only the last result would depend on the relation between the elasticity of the risk-aversion measure and α.[27] However, when the penalty schedule is progressive (with respect to the amount evaded) no clear cut or reasonably simple results emerge with regard to the relations between the incentive to file and the parameters faced by the worker.

So far we have assumed that the number of working hours is fixed by the employer. Allowing the worker to determine his preferred amount of work would induce a decision against filing a return if

$$\max_{H} E\psi(I, T-H) > \max_{H} \psi(\hat{I}, T-H) \tag{19}$$

Denoting by H^* and H^{**} the number of hours that maximizes $\psi(\hat{I}, T-H)$ and $E\psi(I, T-H)$, respectively, the preceding analysis implies that if

$$(\tau-t)wH^*R_A(\hat{I}, T-H^*) < \phi(\rho, \delta) \tag{18'}$$

then

$$E\psi(I, T-H^{**}) \geqslant E\psi(I, T-H^*) > \psi(\hat{I}, T-H^*) \tag{20}$$

Hence, (18') would constitute a sufficient condition for nonfiling when working hours become a choice variable.[28] The interesting question, however, is whether deciding against filing a return involves an increase or a decrease in the amount

---

[27]The left-hand-side of (18) would then be given by $(\tau-t)(w\bar{H})^{\alpha}R_A(\hat{I}, T-H)$.

[28]Since H^* (as well as $T-H^*$) varies (ambiguously) with τ and w, the conclusions derived earlier concerning the effects of these variables on the incentive to file do not apply to this case.

of labor supplied by the worker. In order to inquire into this question suppose now that the utility function is strongly seperable in income and leisure (a frequently used assumption in the tax evasion literature, e.g., Andersen (1977), Pencavel (1979)), so that $\psi(I, T-H) = u(I) + v(T-H)$. By the first-order conditions for maximum utility, H* must then satisfy (expressing $1-\tau$ as $1-t-(\tau-t)$)

$$(1-t)wu'(\hat{I}) - (\tau-t)wu'(\hat{I}) = v'(T-H^*) \tag{21}$$

whereas H** satisfies

$$(1-t)wEu'(I) - \rho\delta(\tau-t)wu'(I^d) = v'(T-H^{**}) \tag{22}$$

We first notice that if the marginal utility of leisure is decreasing, a risk-neutral taxpayer (for whom $u'(\hat{I})=u'(I^d)=Eu'(I)$) will unambiguously supply more hours to work when nonfiling (i.e., $H^{**}>H^*$), since $v'(T-H^{**})$ must be greater than $v'(T-H^*)$ given that $\rho\delta<1$.[29] The intuitive explanation for this result is that nonfiling provides a risk-neutral taxpayer with a higher (expected) net wage rate, $[(1-t)-\rho\delta(\tau-t)]$, which induces him, by a positive substitution effect (there is no income effect as $u''(I)=0$), to increase his work efforts. Should this result extend to risk-aversion depends on whether

$$(1-t)w[Eu'(I)-u'(\hat{I})] > (\tau-t)w[\rho\delta u'(I^d)-u'(\hat{I})] \tag{23}$$

which will definitely hold when $Eu'(I) \geqslant \rho\delta u'(I^d)$. We may thus conclude that if the probability of detection is sufficiently low, i.e., if

$$p \leqslant \frac{1}{1 + (\delta-1)\xi} \tag{24}$$

where $\xi = u'(I^d)/u'(I^{nd})$,[30] a risk-averse taxpayer will increase his supply of

---

[29]As implied by inequality (18') and footnote 26, $\rho\delta<1$ will constitute a sufficient condition for nonfiling under risk-neutrality.

[30]One can show, by differentiating ξ with respect to H, that ξ is constant when relative risk-aversion, $-u''(I)I/u'(I)$, is. The latter (which is satisfied, for instance, for logarithmic utility functions) is a frequently used assumption in analyses of choices involving risk.

labor when deciding against filing (obviously he will do so when $\rho=0$), whereas otherwise his labor supply adjustment remains ambiguous.

Finally, given that the nonfiling condition is satisfied, the effects of marginal parameter changes on the supply of labor are determined by

$$\frac{dH^{**}}{d\tau} = \frac{\delta w \rho u'(I^d)}{\Lambda}[1-R_R(I^d)] \quad (25)$$

$$\frac{dH^{**}}{d\delta} = \frac{(\tau-t)w\rho u'(I^d)}{\Lambda}[1-R_R(I^d)] \quad (26)$$

$$\frac{dH^{**}}{dt} = -\frac{1}{\Lambda}\{-w(1-\rho)u'(I^{nd})[1-R_R(I^{nd})] + w(\delta-1)\rho u'(I^d)[1-R_R(I^d)]\} \quad (27)$$

$$\frac{dH^{**}}{dw} = -\frac{1}{\Lambda}\{(1-t)(1-\rho)u'(I^{nd})[1-R_R(I^{nd})] + [1-t-\delta(\tau-t)]\rho u'(I^d)[1-R_R(I^d)]\} \quad (28)$$

$$\frac{dH^{**}}{d\rho} = -\frac{1}{\Lambda}\{(1-t)w[u'(I^d)-u'(I^{nd})] - \delta(\tau-t)wu'(I^d)\} \quad (29)$$

where $R_R(I)=-Iu''(I)/u'(I)>0$ is the Arrow-Pratt relative risk-aversion measure, and $\Lambda \equiv (1-t)^2w^2(1-\rho)u''(I^{nd}) + [1-t-\delta(\tau-t)]^2w^2\rho u''(I^d) + v''(T-H) < 0$ is the second-order condition for the maximization of expected utility.

Considering these results, notice first that a change in either τ or δ affects the net wage rate in the state of detection only, generating, as expected, opposing income and substitution effects on the supply of labor. Therefore, in analogy to the deterministic model (where $R_R(I)$ is interpreted as the elasticity of marginal utility with respect to income), we obtain that $dH^{**}/d\tau \gtreqless 0$ and $dH^{**}/d\delta \gtreqless 0$ if $R_R(I^d) \lesseqgtr 1$, respectively. However, a change in either t or w affects the net wage rates in both states of the world, requiring an evaluation of the income and substitution effects generated by the change in one rate against those generated by the change in the other. This is possible only if $R_R(I)$ is constant, which implies (after the substitution of the first-order condition in (28)) that $dH^{**}/dw \gtreqless 0$ if $R_R(I) \lesseqgtr 1$, respectively, and that $dH^{**}/dt = 0$ if $R_R(I)=1$, otherwise $\rho \gtreqless 1/[1+(\delta-1)\xi]$

determines that $dH^{**}/dt \gtreqless 0$ if $R_R(I)<1$, but $dH^{**}/dt \lesseqgtr 0$ if $R_R(I)>1$, respectively. A change in ρ, which affects the taxpayer's odds of receiving the alternative net wage rates, would generate opposing substitution effects on the supply of labor, implying that $dH^{**}/d\rho \gtreqless 0$ if $(1-t)/\delta(\tau-t) \gtreqless \xi/(1-\xi)$, respectively. It is interesting to note that in the case of risk-neutrality (where $R_R(I)=0$ and $\xi=1$), a non-filing taxpayer's supply of labor will be negatively related to the tax and law enforcement parameters but positively related to the wage rate, and that in the case of a logarithmic utility function (satisfying $R_R(I)=1$ and $\xi=I^{nd}/I^{d}$), labor supply will be totally insensitive to variations in the taxpayer's environment.

IV. EARNINGS VERSUS WITHHOLDING TAX EVASION

Although the rationale underlying the inception of a withholding system is that taxes become due when incomes are earned (rather than when returns are filed), a supporting belief is that a withholding system helps to assure that taxes due do not escape the tax collector. However, considering the employer's option to substitute for his employees in underreporting their earnings, the ability of a withholding system to favorably affect tax collection is rather questionable. To shed some light on this issue, suppose now that the withholding rate reaches as high as the final tax rate (i.e., that all taxes due are withheld by the employer), that profits are not subjected to taxation and that in the absence of a withholding system a tax-evading wage-earner would face the same law enforcement parameters currently faced by his employer. Following Allingham and Sandmo, a tax-evader's alternative net earnings, Y^{nd} and Y^{d}, would then be given by

$$Y^{nd} = wH - \tau x \tag{30}$$

$$Y^{d} = wH - \tau x - T(wH-x) \tag{31}$$

where $x(\leq wH)$ denotes declared earnings. The expected utility of his prospect would be

$$E\psi(Y, T-H) = (1-p)\psi(Y^{nd}, T-H) + p\psi(Y^{d}, T-H) \tag{32}$$

implying that optimal declaration, x^*, should satisfy

$$\frac{d[E\psi(Y,\ T-H)]}{dx} = -\tau(1-p)\psi_Y(Y^{nd},\ T-H) + (T-\tau)p\psi_Y(Y^d,\ T-H) = 0 \qquad (33)$$

Recalling that under a withholding system the employer chooses Z^* to satisfy

$$\frac{d[EU(\pi)]}{dZ} = -\tau(1-p)U'(\pi^{nd}) + (T-\tau)pU'(\pi^d) = 0 \qquad (34)$$

we obtain that in equilibrium[31]

$$\frac{\psi_Y(Y^{nd},\ T-H)}{\psi_Y(Y^d,\ T-H)} = \frac{U'(\pi^{nd})}{U'(\pi^d)} \qquad (35)$$

That is, the marginal rate of substitution between alternative net earnings, reached - in the absence of a withholding system - by a tax-evading wage-earner, would equal the marginal rate of substitution between alternative net profits, reached - in the presence of a withholding system - by his tax-evading employer.

Evidently, a comparison between the magnitude of evasion in the absence and presence of a withholding system requires assumptions about the exact shape of the employer's and his employees' utility functions. Let us adopt the frequently applied logarithmic function for both $\psi(Y,\ T-H)$ and $U(\pi)$,[32] assuming also that all workers employed by given employer have identical tastes. Equation (35) is then reduced into

$$\frac{Y^d}{Y^{nd}} = \frac{\pi^d}{\pi^{nd}} \qquad (36)$$

which, after multiplying Y^{nd} and Y^d by the number of employees, N/H, substituting and rearranging, becomes

$$\frac{\tau(wN-X^*)}{\tau(wN-Z^*)} = \frac{(1-\tau)wN}{V(N)-wN} \qquad (37)$$

---

[31]Notice that $\tau > pT$ is sufficient to assure that either $x^* < wH$ or $Z^* < wN$.

[32]For example, Isachsen and Strom (1980) used a strongly seperable logarithmic utility function ($\psi = \ln Y + \ln(T-H)$) to facilitate their analysis of the individual's choice between registered and unregistered work. <u>Strong</u> seperability is not required for ... discussion, however.

where $X(=xN/H)$ denotes earnings declared by all employees in the absence of a withholding system. We thus conclude that if the employer's and workers' preferences can be approximated by logarithmic functions, the ratio between the amount of tax evaded in the absence of a withholding system and the amount escaping the tax collector in the presence of a withholding system would equal the ratio between legitimate net earnings and legitimate net profits. *Hence, a withholding system would help to reduce evasion* ($Z^*>X^*$) only if the workers' legitimate share in output exceeds the legitimate share of their employer.[33] Obviously, a withholding system would not guarantee a reduction in the amount evaded under more general specifications of the employer's and employees' preferences.

Given the operation of a withholding system we may also question how a reduction in the withholding tax rate below the final rate affects the total amount evaded by the employer and his workers. Beginning with a moderate reduction in the withholding rate which preserves the nonfiling condition, the total amount escaping the tax collector, S, will be

$$S = t(wN-Z^*) + (\tau-t)wN = \tau wN - tZ^* \qquad (38)$$

as compared to $\tau(wN-Z^*)$ when $t=\tau$. Recall now that Z^* is negatively or positively related to t, depending on whether the penalty multiplier is fixed ($T=\bar{T}$) or varies with t ($T=\lambda t$), respectively (Table 1). Hence, a (small) reduction in the withholding rate below the final rate will have an undeterminate effect on evasion under the former penalty scheme, while unambiguously increasing evasion under the latter. As one continues to lower the withholding tax rate, evasion will continue to increase

[33]Alternatively, rewriting (36) as $[Y^{nd} - T(wN-X^*)]/[\pi^{nd} - T(wN-Z^*)] = Y^{nd}/\pi^{nd}$, yields first

$$\frac{wN-X^*}{wN-Z^*} = \frac{Y^{nd}}{\pi^{nd}} \qquad (37')$$

which implies that $Z^*>X^*$ if $Y^{nd}>\pi^{nd}$ (or, by (36), if $Y^d>\pi^d$). However, since Y^{nd} and π^{nd} vary with X^* and Z^*, we may express Y^{nd} and π^{nd} as $(1-\tau)wN + \tau(wN-X^*)$ and $V(N)-wN + \tau(wN-Z^*)$, respectively, and substitute into (37') to obtain (37).

under the latter scheme, until the withholding rate is sufficiently reduced to induce filing.[34] Hereafter, the amount evaded becomes $t(wN-Z^*)$, on which further reductions in the withholding rate will have an indeterminate effect. At this point, however, the alternative penalty scheme contributes an unambiguous result, as further reductions in the withholding rate will clearly help to reduce the amount evaded.

V. SUMMARY AND CONCLUDING REMARKS

We have discussed the employer and the employee tax fraud activity under a withholding tax system. After identifying the conditions sufficient to induce a risk-averse employer to avoid full-remittance of his employees' withheld taxes by deliberately understating his actual wage payments, and after establishing the neutrality of these conditions with regard to his employment decision, we have concentrated on examining the relations between fraudulent declaration of wage payments, the withholding and profit tax rates and the competitive wage rate under alternative penalty schemes. In the absence of profit taxation, an increase in the withholding tax rate has been found to have an unambiguous negative effect on declaration when the penalty relates (proportionally) to the level of unstated wage payments, but an unambiguous positive effect on declaration when the penalty relates to the amount of non-remitted taxes. While the presence of profit taxation would not affect the former conclusion regardless of tax authorities' treatment of profit tax overpayments occurring due to the understatement of wage payments, it would retain the latter conclusion if the penalty is assessed on the <u>net</u> amount evaded, but would lead to ambiguity if profit tax overpayments are simply refunded upon detection or are totally disregarded by the tax authorities. An increase in the profit tax rate would

---

[34]Assuming, for example, that all workers' utility functions are logarithmic, the nonfiling condition will be $(\tau-t)/(1-\tau) < \phi(\rho, \delta)$. Hence, they will all decide in favor of filing when t is reduced below $\tau-(1-\tau)\phi(\rho, \delta)$.

have, however, an unambiguous positive effect on declaration - regardless of whether the penalty relates to understating or to non-remittance - if the employer is not credited in any way for his profit tax overpayments, a negative effect on declaration when the penalty is assessed on the net amount evaded and given that the penalty multiplier exceeds or equals the ratio between declared profits and unstated wages - and an ambiguous effect otherwise (including the case where overpayments are just refundded upon detection). Finally, the effect on declaration of an increase in the competitive wage rate has been found, taking account of its effect on employment, to depend upon the elasticity of labor demand, being positive for all penalty schemes when the latter is less than or equals unity, negative when it equals or exceeds a varying (with the penalty scheme) critical value and ambiguous otherwise.

Addressing the employee's decision on whether to evade his non-withheld taxes by avoiding filing a final tax return has revealed, as intuitively expected, that the incentive to file - when the number of working hours is fixed by the employer - will be lower the lower is the level of enforcement or the lower is risk-aversion at the state of filing, but contrary to intuition, the lower is the final tax rate and the closer to the latter is the withholding tax rate. Also, the incentive to file has been found to be lower the lower or the higher is the wage rate if the elasticity of risk-aversion with respect to legitimate net earnings is smaller or greater than unity, respectively. Allowing the worker to determine his preferred amount of work, we have been able to show - given strong seperability of the utility function - that deciding against filing a return will be accompanied by increased supply of a risk-neutral's labor, and - if the probability of detection is sufficiently low - of a risk avoider's as well. Given nonfiling, a risk-neutral's supply of labor will be negatively related to the tax and law enforcement parameters but positively related to the wage rate, whereas a risk-avoider's supply of labor may exhibit all possible responses to parameter variations, depending upon the magnitudes of his relative risk-aversion and marginal rate of substitution between net earnings in the states of detection and non-detection.

Comparing between the magnitude of tax evasion in the absence and presence of a withholding system has revealed that a withholding system can not guarantee a reduction in the amount evaded. Using logarithmic utility functions we have found that in the absence of profit taxation, a withholding system which allows an employer to withhold the entire amount of his employees' tax liability would help to reduce evasion only if their legitimate share in output exceeds his own. Given the existence of a withholding system, a moderate reduction in the withholding tax rate below the final rate (which is sufficient to induce evasion of non-withheld taxes) would have, regardless of the exact shape of the utility functions, an ambiguous effect on the total (withheld and non-withheld) amount evaded if the penalty imposed on a detected employer relates to unstated wage payments, but would unambiguously increase the amount evaded if the penalty relates to non-remitted taxes. As the withholding rate is pushed further down, the amount evaded would continue to increase - given the latter scheme - until the withholding rate has been sufficiently reduced to eliminate the incentive for nonfiling. Hereafter, a reduction in the withholding rate would have an ambiguous effect on the amount evaded, but would unambiguously succeed in reducing it under the alternative penalty scheme.

Being widely accepted as a tax collection mechanism, it is remarkable that the withholding system has not been subject to any analytical treatment in the public finance literature. While having the advantage of linking tax payments to the current level of income (rather than having them lag behind for one year), the withholding system has its operation costs. One may question whether the system is economically justified or how withholding costs affect employers' demand for labor. Given that the system is desirable, a tax planner must decide whether all tax liability should be deducted at source, allowing the exemption of single job holders with no other income from the obligation of filing a return (the Israeli case) - or whether the withholding rates should not reach as high as the final rates, requiring all wage-earners to file a return (the United States' case). Adopting the latter scheme raises the problem of determining the withholding tax rates. As implied by this paper, tax evasion considerations must be incorporated into any attempt to establish a tax collection system which is socially optimal.

REFERENCES

Allingham, M.G. and A. Sandmo, 1972, Income tax evasion: A theoretical analysis, Journal of Public Economics 1, 323-338.

Andersen, P., 1977, Tax evasion and labour supply, Scandinavian Journal of Economics 79, 375-383.

Baldry, J., 1979, Tax evasion and labour supply, Economics letters 3, 53-56.

Cowell, F.A., 1985, Tax evasion with labour income, Journal of Public Economics 26, 19-34.

Isachsen, A.J. and S. Storm, 1980, The hidden economy: The labor market and tax evasion, Scandinavian Journal of Economics 82, 304-311.

Koskela, E., 1983, A note on progression, penalty schemes and tax evasion, Journal of Public Economics 22, 127-133.

Marrelli, M., 1984, On indirect tax evasion, Journal of labor Economics 25, 181-196.

Peckman, J.A., 1983, Federal tax policy (The Brookings Institute, Washington D.C.)

Pencavel, J.H., 1979, A note on income tax evasion, labor supply, and nonlinear tax schedules, Journal of Public Economics 12, 115-124.

Weiss, L., 1976, The desirability of cheating incentives and randomness in an optimal income tax, Journal of Political Economy 84, 1343-1352.

Yaniv, G., 1988, A note on the interpretation of the income effect in tax evasion models (in preparation).

Yitzhaki, S., 1974, A note on income tax evasion: A theoretical analysis, Journal of Public Economics 3, 201-202.

NATIONAL INSURANCE INSTITUTE

Bureau of Research and Planning

Jerusalem, March 1988

Discussion paper 42